Searching for supersymmetry in the first LHC data with ATLAS

Searching for supersymmetry in the first LHC data with ATLAS

Een wetenschappelijke proeve op het gebied van
Natuurwetenschappen, Wiskunde en Informatica

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op maandag 23 januari 2012
om 10.30 uur precies

door

Folkert Koetsveld

geboren op 10 januari 1979
te 's Gravenhage

Promotores:	Prof. dr. N. de Groot	
	Prof. dr. ir. P.J. de Jong	
Manuscriptcommissie:	Prof. dr. M.H.M. Merk	Vrije Universiteit
	Prof. dr. R.H.P. Kleiss	Radboud Universiteit
	Dr. A.C. König	Radboud Universiteit

The work described in this thesis is part of a Projectruimte “SUSY with muons”. The author was financially supported by the Foundation for Fundamental Research on Matter (FOM).

Cover illustration: De Oerknal, painting by Jaap Stomphorst. Design by Michiel Stomphorst (http://www.ideaalcommunicatie.nl).

Contents

Introduction

Since physicists have started the design of the LHC accelerator and the ATLAS detector, many predictions of the Standard Model (SM) have been tested and confirmed with very high accuracy. Open questions of the SM today are the masses and mixings of neutrinos, and the status of the Higgs mechanism for electroweak symmetry breaking. Answering the latter of these questions is one of the goals of the ATLAS experiment. Further unanswered questions deal with the existence of physics beyond the SM, and form a second motivation for the LHC and for ATLAS.

Cosmological observations have made it clear that the SM is not enough to describe all phenomena in the observable universe. First, the accelerated expansion of the universe indicates the existence of an unknown form of dark energy. Second, evidence for the existence of cold dark matter can be found in observations of rotational velocities of stars around a galaxy, of gravitation lensing effects near colliding galaxies, and theoretically from models of clustering of galaxies. There are no particles in the SM that can be used to create cold dark matter with the required density. Therefore a new particle is predicted, which may be (in-)visible in the proton-proton collisions of the LHC.

One possible extension of the SM which contains particles that could be the constituent of dark matter has been around for a long time. Originally, supersymmetry, or SUSY, was devised in the 1970's as an extension to the symmetries of nature and as a part of string theory. It has fallen in and out of favor over the years, and several times has been proclaimed dead or dying. However, when it was shown that SUSY with the added symmetry of R-parity conservation (the conservation of SUSY-number) is capable of producing a stable weakly interacting massive particle with the correct cosmic density, many experimenters regained interest. Nowadays SUSY is used as a benchmark for new physics in the community of particle physics experimentalists. To this end extensive modeling and simulation of the phase space of SUSY parameters has been done.

The first chapter of this thesis will discuss both the SM and SUSY in some detail, specifically describing experimental signatures of both theories, which may be exploited in the search for physics beyond the SM presented here.

Finding evidence for a new theory without knowing exactly what the experimental result might look like poses a challenge on both the accelerator and the detector. Since the energy scale of the new theory is unknown, the highest possible energy needed to be reached. The Large Hadron Collider was designed to collide protons at a center of mass energy of 14 TeV. Since the collisions producing new particles could be very rare, an instantaneous luminosity up to $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ was required.

As a candidate Dark Matter particle is expected to be subject only to the weak interaction, it will not leave a signal in the detector. Therefore an excellent coverage in all directions was required, in order to use the missing transverse momentum in collisions as evidence for these invisible particles. Many more requirements were made on both detector and accelerator, which will be discussed in detail in the second chapter of this thesis. With the recording of the first data in 2009 and 2010, the actual performance of the detector has been evaluated, and the result will also be discussed in the second chapter.

When looking for evidence of new physics, it is of great importance that the known physical theory is well understood. The SM of particle physics has never been tested at the energies delivered by the LHC. Theoretical predictions of the behavior of the SM are therefore extrapolations of low-energy behavior. Before any claim of new phenomena can be made, these extrapolations have to be tested and understood. As long as these tests are not completed, analyses of high energy data must not rely on the information of current Monte Carlo (MC) generators, which for now are based on these extrapolations. Chapters 6 and 7 describe the data driven combined fit method, which aims to be as independent of MC information as possible.

In 2010 the ATLAS detector recorded the first pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. The final size of the recorded sample was $L_{int} = 35\,\mathrm{pb}^{-1}$. This dataset already made the ATLAS collaboration sensitive to parts of the SUSY phase space that had not been probed before. Evidence of new physics could be present in this dataset, and if no evidence is found, a new part of the SUSY phase space can be excluded. The last chapter of this thesis describes the result of applying the combined fit method on the 2010 ATLAS data.

Chapter 1

The Standard Model and supersymmetry

1.1 The Standard Model

The theory which is called the Standard Model (SM) of high energy physics aims to describe all matter found in nature and all interactions except gravity. This description agrees with the experimental results with unprecedented accuracy.

The elementary particles that are the building blocks of all known matter are fermions, spin-$\frac{1}{2}$ particles. These fermions make up protons, neutrons and atoms. They can be divided into two groups by their interactions: leptons, which undergo the electro-weak interaction, and quarks, which couple to both the electro-weak and the strong interaction. The quarks and leptons that form 'ordinary' matter, i.e. the elements we know from chemistry, are the electron, the up and the down quark. Together with the electrically neutral electron-neutrino, which shows up in e.g. nuclear decay of unstable atoms, they form the first of three families of matter particles that are known.

Each particle in the first family has a corresponding particle with different mass in the each of the other two families. The mass of the quarks and charged leptons increases from the first to the second and from the second to the third family. Although (at least two) neutrinos are massive, their mass ordering is unknown.

Apart from the fermions that make up matter, the SM contains spin-1 bosons that are the mediators of the three known interactions. The first is the photon, which mediates the electromagnetic interaction. The $W^{\pm}$ and Z bosons govern the weak interaction, seen in nature when atomic nuclei undergo radioactive decay. Finally the gluons (eight in total) are responsible through the strong force for binding quarks into protons and other hadrons, and (indirectly) the formation of atomic nuclei. The only unobserved particle that is a part of the Standard Model is the Higgs particle, which is the quantum of the field that is hypothesized to break the electroweak symmetry in the SM, and give particles their masses. All particles in the Standard Model (except for the Higgs) are listed with some of their properties in Figure 1.1.

The Standard Model is described in the language of quantum field theory (QFT), where all particles and interactions are described by a field in spacetime. The Lagrangian of the SM

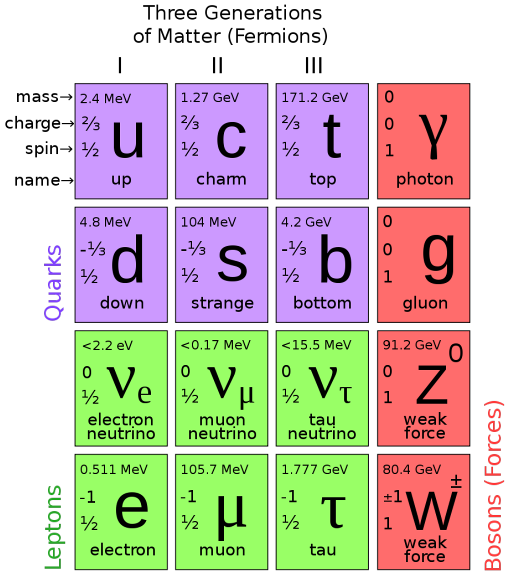


Figure 1.1: Schematic of all particles in the Standard Model (except the Higgs boson).

specifies the mass and interactions for all known particles. The first part of this chapter will look at different aspects of the Standard Model, starting from the basics of QFT, and the use of symmetry groups in describing interactions. The next section handles some of the details of the electroweak sector. Coupled to the electroweak sector the Higgs mechanism of spontaneous symmetry breaking will be discussed. The first part of this chapter ends with a description of the strong interaction, the structure of the proton and the unique signature of proton-proton collisions.

The second part of this chapter is a description of a possible extension of the SM, namely supersymmetry. This extension is the subject of the research in this thesis. Starting from a minimal supersymmetric model with only one set of particles, this model will be extended to include interactions and the supersymmetric partners of all SM particles. Then the mechanism of supersymmetry breaking will be discussed, ending with a description of mSUGRA, a specific model of broken supersymmetry.

1.1.1 Simple Lagrangian

The simplest Lagrangian for a free, massless, fermion field $\psi(x)$ contains just a kinetic term:

$$\mathcal{L}_{ff} = \bar{\psi} i \gamma^{\mu} \partial_{\mu} \psi, \tag{1.1}$$

with γ^{μ} the Dirac matrices. To this simple Lagrangian interactions can be added in the form of gauge symmetries. There are three types of interactions in the SM (not counting the Higgs sector for now), which are the electromagnetic, the weak and the strong interaction. Within the formalism of the SM the first two forces are combined in the electroweak interaction, see section 1.1.2. Each interaction is the result of a symmetry present in the Lagrangian of the

SM. For instance, the Lagrangian of the free fermion in equation 1.1 is invariant under the gauge transformation:

$$\psi(x) \rightarrow e^{i\alpha}\psi(x) \tag{1.2}$$

Just like translation and rotation are transformations coupled to a symmetry in spacetime, equation 1.2 is coupled to a symmetry of the free fermion theory in the complex plane. The symmetry group to which this transformation belongs is the unitary group U(1). One can think of other symmetries that leave the free fermion Lagrangian intact. For instance, the field $\psi(x)$ can belong to a representation of $SU(2)$ or $SU(3)$, the unitary rotation groups. Generators of these groups induce rotations. Such a transformation looks like:

$$\psi(x) \rightarrow e^{iV^a\beta^a}\psi(x)_i \tag{1.3}$$

with the index a running over all the generators of the the gauge group, and the index i running over the dimensions on which the generators act.

These global symmetries are present in the free fermion Lagrangian 1.1. Interactions can enter the Lagrangian when postulating that these gauge symmetries are local symmetries of the theory, i.e. a different gauge transformation can be chosen for each point in space-time. According to Noethers theorem, these gauge symmetries are observed in nature as conservation laws. The $U(1)$ symmetry of electromagnetism yields the conservation of electric charge. Conservation of lepton and baryon number indicates the (approximate) $SU(2)$ symmetry of the weak interaction, and the color charge of the quarks inside the hadrons is a result of the $SU(3)$ symmetry of the strong interaction.

By imposing locality, a gauge boson is introduced for each generator of the corresponding gauge group. These particles are the photon, the W and Z bosons and the gluons. Their place in the SM will be discussed in the next sections.

How local gauge invariance introduces new fields can be shown using a simple quantum field theory, pure Quantum ElectroDynamics (QED, the first quantum field theory describing electromagnetic interactions). The local gauge transformation can be written as:

$$\psi(x) \rightarrow e^{i\alpha(x)}\psi(x) \tag{1.4}$$

When performing this transformation on the fields in the Lagrangian of equation 1.1, one gets:

$$\bar{\psi}i\gamma^\mu\partial_\mu\psi \rightarrow \bar{\psi}e^{-i\alpha(x)}i\gamma^\mu(e^{i\alpha(x)}\partial_\mu\psi + ie^{i\alpha(x)}\psi\partial_\mu\alpha(x)) = \bar{\psi}i\gamma^\mu(\partial_\mu\psi + i\psi\partial_\mu\alpha(x)) \tag{1.5}$$

Hence the Lagrangian is not invariant under the local symmetry. To restore the invariance of the theory to this gauge transformation one can replace the derivative with the so-called covariant derivative:

$$D_\mu \equiv (\partial_\mu + ieA_\mu) \tag{1.6}$$

where the vector potential A_μ transforms like:

$$A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu\alpha(x) \tag{1.7}$$

Inserting the covariant derivative 1.6 in the Lagrangian 1.1 gives:

$$\mathcal{L}_{QED} = \bar{\psi}i\gamma^\mu D_\mu\psi = \bar{\psi}i\gamma^\mu\partial_\mu\psi - e\bar{\psi}\gamma^\mu A_\mu\psi \tag{1.8}$$

The first term on the right hand side of this equation shows again the original kinetic term of the free fermion. The second term shows an interaction with the new field, with coupling strength e. When transforming this Lagrangian under the local $U(1)$ symmetry, the $\partial_\mu \alpha(x)$ terms cancel and the invariance of the theory under the local transformation is restored.

Thus imposing a local gauge symmetry on the model leads to the introduction of a new field, which interacts with the original fermionic fields. The $U(1)$ group describes abelian symmetry, meaning that the operators of the group commute. In contrast, the $SU(2)$ and $SU(3)$ groups are non-abelian. The principle of introducing a covariant derivative remains the same. For instance the transformation of a fermion field under $SU(2)$ would be given by equation 1.3 with $V_a(x)$ a set of three matrices that are generators of $SU(2)$ (the Pauli matrices for instance). Inserting this in the free fermion Lagrangian yields

$$\mathcal{L}_{SU(2)} = \bar{\psi} i \gamma^\mu (D_\mu) \psi = \bar{\psi} i \gamma^\mu \partial_\mu \psi - g \bar{\psi} \gamma^\mu A^a_\mu V^a \psi \tag{1.9}$$

introducing three new vector fields A^a_μ, one for each generator of the $SU(2)$ group, and a coupling strength g.

The strong interaction Lagrangian is invariant under local $SU(3)$ transformations. It looks similar to equation 1.9, but now the ψ_i has three components, and there are 8 generators of the group, hence there are eight gauge fields V^a_μ. In the next sections the electroweak and strong sectors will be discussed in more detail.

All particles in the theory so far are massless. However, the particles that make up matter are known to be massive. Also, the weak force, which can be derived by imposing a local $SU(2)$ symmetry on the model, has a limited range, caused by the mass of its mediator bosons. In the Lagrangian of equation 1.9 however, a mass term for the bosons would break the gauge symmetry. Adding a mass term for a fermion will also break gauge symmetry, as will be demonstrated later in this chapter.

The Higgs mechanism will provide a way out, but this only works when electromagnetism and the weak force are two (low-energy) branches of the same (high-energy) tree of the electroweak force, which has the structure of the combined symmetry group $SU(2) \times U(1)$. The electroweak sector of the SM will be described in the next section.

1.1.2 Electroweak interactions

The theory of the electroweak interactions proposed by Glashow, Weinberg and Salam (GWS theory) [1, 2, 3] takes as a starting point a $U(1) \times SU(2)$ gauge symmetry. It was proposed as a unification of the weak and the electromagnetic interactions. In unbroken form, all its gauge bosons are massless, as required by gauge invariance. The charge of the $U(1)$ symmetry is called hypercharge, and it will turn out to be different from the electric charge of the broken EW model. After breaking through the Higgs mechanism (which will be described in the next section) the gauge bosons mix to form three massive particles ($W^\pm$ and Z) for the weak force and one massless particle, the photon of electromagnetism.

In the light of the weak interaction it is necessary to discuss the chirality of the fermions of the Standard Model. A fermion ψ in the Dirac representation is a superposition of two chirality eigenstates ψ_L and ψ_R:

$$\psi_L = (1 - \gamma^5)\psi \quad \text{and} \quad \psi_R = (1 + \gamma^5)\psi. \tag{1.10}$$

The representation where fermions are chirality eigenstates is called the Weyl representation. This representation will be used extensively when describing the supersymmetric extension of the SM later in this chapter.

The weak force has the peculiar property that it only couples to fermions of left-handed chirality. It combines left-handed fermions in doublets, where the electric charge of the components of one doublet differs by 1 unit. In the SM, the electron is part of a doublet with the electron neutrino, and the up-quark is the weak partner of the down quark, see also Figure 1.1: each quark in the top row is coupled to the quark directly below, and each neutrino in the third row is coupled to the corresponding charged lepton in the fourth row.

Right-handed fermions are $SU(2)$ singlets. They may still couple to the electromagnetic (charged leptons and quarks) and the strong (quarks) force. Returning now to the mass terms for the fermions of the SM, the problem of these terms can now be illustrated using the Weyl representation. A fermion mass term, which mixes left- and right-handed states:

$$m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \tag{1.11}$$

contains terms that explicitly break the gauge invariance of the Lagrangian since the left- and right-handed fields have different weak charges. Also the $U(1)$ hypercharge for these components is different (-1 for ψ_R, $-\frac{1}{2}$ for ψ_L). However, as we see masses of particles in nature the symmetry must be broken. The mechanism that is mostly used to introduce mass in the Standard Model is the Higgs mechanism, which will be discussed in the next section.

1.1.3 The Higgs mechanism

From the previous section it is clear that some method of symmetry breaking needs to be introduced in the SM. In order to save the properties of local gauge invariance that make the theory so successful, a method has been introduced that breaks the symmetry spontaneously. Spontaneous symmetry breaking leaves the symmetry of the Lagrangian intact, but allows for ground states which are asymmetric. The Higgs mechanism spontaneously breaks the $SU(2) \times U(1)$ gauge symmetry of the electroweak interaction (see next section), and in the process generates a mass term for the charged fermions and the bosons of the weak interaction. The photon and the gluons remain massless. The mass of neutrinos is a more complex story, very interesting in itself but not relevant for this thesis. This subject will not be covered here. The Higgs mechanism was introduced by Goldstone [4], developed independently by Higgs [5], Guralnik, Hagen and Kibble [6] and Englert and Brout [7] and introduced into the SM by Glashow, Weinberg and Salam.

Before turning to the Higgs mechanism as it appears in the electroweak theory, I present a toy model with an electrically charged spin-0 field added to the QED model. The Lagrangian for such a model would be:

$$\mathcal{L} = -\frac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 + \mu^2\phi^*\phi - \frac{\lambda}{2}(\phi^*\phi)^2 \tag{1.12}$$

with D_μ the covariant derivative of the electromagnetic interaction (equation 1.6), and $F_{\mu\nu}$:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.13}$$

the kinetic term of the gauge field A_μ. If the parameter μ^2 is negative the field ϕ can acquire a vacuum expectation value (v.e.v.) due to the shape of the potential, see Figure 1.2. In

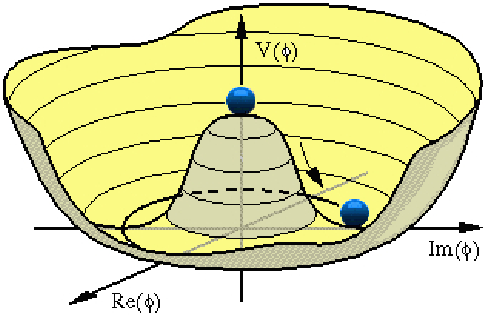


Figure 1.2: The "Mexican Hat" Higgs potential in 2 dimensions

this figure one sees that the potential is symmetric under rotations. When the field enters its ground state however, it must choose a direction since the minimum is away from the origin. Choosing any direction obviously breaks the symmetry of Figure 1.2. One can now redefine the field ϕ by expanding it around its v.e.v.:

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}}\left(\phi_1 x + i\phi_2(x)\right) \tag{1.14}$$

Through the coupling of $\phi(x)$ to the field A_μ in the covariant derivative, one gets a mass term for the field A_μ:

$$e^2\phi_0^2 A_\mu A^\mu = \frac{1}{2} m_A^2 A_\mu A^\mu \tag{1.15}$$

In this way, a mass term can be added the Lagrangian in equation 1.12 without breaking gauge symmetry explicitly. It is said that the symmetry is spontaneously broken.

1.1.4 Electroweak symmetry breaking

Now that the Higgs mechanism is defined, this section will show how it is applied to the $(SU(2) \times U(1))$ electroweak theory. The scalar part of the electroweak Lagrangian is:

$$\mathcal{L} = |D_\mu \phi|^2 + \mu^2 \phi^* \phi - \frac{\lambda}{2}(\phi^* \phi)^2 \tag{1.16}$$

but now the covariant derivative working on the scalar field will contain the gauge fields of the EW interaction, A_μ^a and B_μ for $SU(2)$ and $U(1)$ respectively:

$$D_\mu \phi = (\partial_\mu - \frac{1}{2} ig A_\mu^a \sigma^a - i\frac{1}{2} g' B_\mu)\phi \tag{1.17}$$

where g and g' the respective coupling strength. The field ϕ is now a $SU(2)$ doublet:

$$\phi(x) = \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix}. \tag{1.18}$$

with hypercharge $\frac{1}{2}$. A gauge transformation on a scalar field in this theory would look like this:

$$\phi \rightarrow e^{i\alpha^a \sigma^a/2} e^{i\beta/2} \phi \tag{1.19}$$

with $\sigma^a, a \in \{1,2,3\}$ the Pauli matrices, which are generators of the $SU(2)$ group. One can use the rotational freedom of $SU(2)$ to make sure that the vacuum expectation value lies in a specified direction, like this:

$$< \phi >= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.20}$$

where v can be made real through a $U(1)$ transformation. Evaluating the $|D_\mu \phi|^2$ term of equation 1.16 at the v.e.v. of the scalar field gives:

$$(D_\mu \phi_0)^2 = \frac{1}{2}\frac{v^2}{4}[g^2(A^1_\mu)^2 + g^2(A^2_\mu)^2 + (g' B_\mu - g A^3_\mu)^2] \tag{1.21}$$

Here one sees the mass terms appear. Diagonalizing the mass matrix yields four gauge bosons, three with mass (W^+, W^- and Z, and one massless (the photon). In terms of the A^a_μ and B_μ fields, these are:

$$\begin{aligned} W^\pm_\mu &= \frac{1}{\sqrt{2}}(A^1_\mu \mp i A^2_\mu); & m_W &= \frac{gv}{2} \\ Z_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g A^3_\mu - g' B_\mu); & m_Z &= \sqrt{g^2+g'^2}\frac{v}{2} \\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g A^3_\mu + g' B_\mu); & m_A &= 0 \end{aligned} \tag{1.22}$$

From the W boson mass one can derive that the v.e.v. of the Higgs field should be 246 GeV. Unfortunately this does not say anything about the actual mass of the Higgs. Fits to electroweak data that do give an indication of the mass of the Higgs boson will be discussed later.

Fermions also obtain mass via the Higgs mechanism. As said, one cannot carelessly enter a fermion mass term in the SM Lagrangian. The reason is that only left-handed fermion doublets couple to the A^a_μ fields of $SU(2)$. The right-handed fermions are singlets under SU(2). Considering the mass term of equation 1.11, it is clear that this term is not invariant under SU(2), and the hypercharges of the mass terms do not add up to 0.

Again the Higgs mechanism has a way out. By coupling the left-handed $SU(2)$ doublet field E_L containing the left-handed electron and electron neutrino to the Higgs doublet field and the right-handed electron one gets:

$$-y_e \bar{E}_L \phi e_R + h.c. \tag{1.23}$$

with y_e the strength of the Yukawa coupling between the electron and the Higgs. The left-handed doublet and the Higgs form a $SU(2)$ singlet. The Higgs doublet has hypercharge $\frac{1}{2}$, so the hypercharges of the term in equation 1.23 add up to 0. This term leaves the gauge symmetry invariant. When replacing the Higgs field with the vacuum expectation value given in equation 1.20, one gets a mass term coupling the left-handed electron to its right-handed counterpart:

$$\frac{1}{\sqrt{2}} y_e v \bar{e}_L e_R + h.c. \tag{1.24}$$

while the $U(1) \times SU(2)$ symmetry of the Lagrangian remains intact.

1.1.5 Strong interaction

The strong interaction is relatively simple in its formal Lagrangian, and very complicated in experimental practice. The symmetry group of the strong interaction is the non-abelian group $SU(3)$, which contains eight generators, corresponding to eight gluons, the mediator bosons of the strong interaction. The charge of the strong interaction is called color charge. It has three values, commonly called red, green and blue. A combined state of the three different colors is color-neutral, or 'white'. The only fermions in the SM that couple to color are the six quarks. So each left-handed quark has an electric charge, is part of a $SU(2)$ doublet and also of a $SU(3)$ color triplet. For more information about the strong interaction see the book by Ellis, Stirling and Webber [8] and the book by Peskin and Schroeder [9].

There are two peculiar properties of the strong interaction which make it challenging to work with both theoretically and experimentally: confinement and asymptotic freedom. Asymptotic freedom relates to the running of the strong coupling. Due to vacuum polarization effects, all couplings run, i.e. have a value which depends on the energy level at which they are evaluated. The electromagnetic coupling is small at low energies, and increases with increasing energy. This can be understood intuitively, as the polarization of the vacuum around a charged particle screens it from its surroundings, reducing its 'visible' charge. At higher collision energies this screening is less effective, and with increasing energy more of the naked electric charge of a particle is visible. The strong coupling however is large at low energies. In fact it is so large that perturbation theory, the method used for most quantum field calculations, is not applicable. At larger energies the coupling becomes smaller, which means that the polarization enhances the strength of the coupling instead of screening it. Thus the strongly interacting particles behave more like free particles at higher energy, hence the term asymptotic freedom.

The other peculiar property of the strong interaction is that of confinement, the fact that no naked color charge has ever been observed. Quarks are always in a bound state of three quarks (baryons) or a quark and an antiquark (mesons).

Together, confinement and asymptotic freedom (coupled to the large strength of the strong interaction at low energies) give rise to two features of strong interaction processes that are important when studying physics in a collider, especially a hadron collider such as the LHC. One is the structure of the proton, and the other is the formation of jets.

Formation of jets

To illustrate the formation of jets, it can be useful to look at the simplest QCD process, which is an quark anti-quark pair annihilating into a gluon, which then decays into two quarks. At lowest order, this process can be described by the Feynman diagram of Figure 1.3 (a). Looking one order higher in α_s, one gets contributions like Figure 1.3 (b). The added gluon on outgoing quark lines is commonly called final state radiation (FSR), or initial state radiation (ISR) if the gluon is radiated from the initial pair of quarks. This diagram has a large contribution when the radiated gluon is collinear, or relatively soft. In fact any diagram with any number of soft or collinear radiated gluon contributes significantly to the total cross section.

This is taken into account using so-called parton shower models. Starting from the final state of the matrix element calculation (called the hard interaction) these models will add soft and collinear radiation to calculate a final state of naked color particles.

Since single color charged particles are not observed in nature, some form of hadronization will take place. This is a soft process, at energies where α_s becomes too large to perform

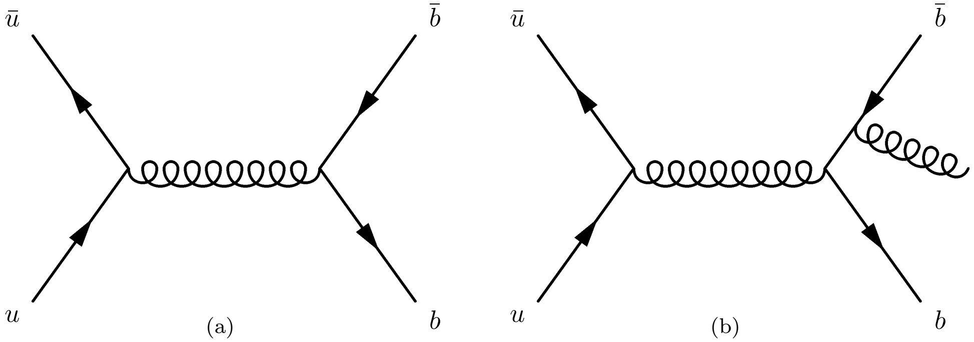


Figure 1.3: Feynman diagram for the process $u\bar{u} \rightarrow b\bar{b}$ (a) at highest order in α_s and (b) one of the terms at next-to-leading order in α_s.

perturbative calculations. Because showering and hadronization take place at lower energies, and thus at longer timescales, it is deemed justified to treat the hard interaction, subsequent parton showering and hadronization as separate processes. This separation of any QCD process into a hard and a soft part is called factorization.

After the parton shower and hadronization, one ends up with a jet of particles, which enter the detector. The definition of a jet, both experimental and theoretical, is non-trivial. There are many definitions available [10]. An important qualification of a jet definition is infrared and collinear safety, which means that a collinear or very soft gluon that is radiated from any colored particle in the interaction should not change the outcome of a jet finding algorithm. For a long time, the definitions used in experiment did not have this property, but lately there has been a lot of improvement in this area. In fact, the jet reconstruction algorithm used in chapter 6, which is a simple cone algorithm, is not infrared and collinear safe. When working on first data a different, better algorithm was used. These algorithms will be described in more detail in the next chapter.

Structure of the proton

The proton is not a point-like particle. In the simplest picture it is a color-neutral bound state of three colored quarks, called the valence quarks. This simple picture was tested and disproved by Deep Inelastic Scattering (DIS) experiments, where an electron (or a neutrino) is scattered off a proton (or neutron). The proton is shattered by the collision, hence the name inelastic scattering.

Taking an collision between a proton of momentum p and mass M, and an incoming electron, where q is the momentum of the virtual photon that is exchanged between the two particles, the cross section of this process can be written as:

$$\frac{d\sigma}{dE' d\Omega} = \left(\frac{\alpha}{2E\sin^2(\theta/2)}\right)\left[2W_1(q^2,x)\sin^2(\theta/2) + W_2(q^2,x)\cos^2(\theta/2)\right] \tag{1.25}$$

with θ the scattering angle of the electron, E and E' the initial and final electron energy and

$$x = -\frac{q^2}{2q\cdot p}. \tag{1.26}$$

The $W_1(q,x)$ and $W_2(q,x)$ are form factors that parametrize the internal structure of the proton. For point-like constituents, colliding at higher energy will not reveal new structure of the proton, and these form factors should be independent of the energy transfer q. This effect, known as Bjorken scaling, was confirmed by experiment at energies larger than about 1 GeV. This means that W_1 and W_2 in equation 1.25 get replaced by:

$$MW_1(q^2, x) \rightarrow F_1(x) \tag{1.27}$$

$$-\frac{q^2}{2Mx} W_2(q^2, x) \rightarrow F_2(x) \tag{1.28}$$

It shows that an electron scattering of a proton sees only one of the quark constituents. The softer interaction between the quarks happens over much longer time scale than the DIS process, and does not influence the cross section, meaning that the quark is essentially free. Thus the process is reduced to the scattering of two point-like particles. This can be used to get an exact form of the structure functions:

$$F_2(x) = 2xF_1(x) \tag{1.29}$$

$$F_1(x) = \frac{1}{2} \sum_i Q_i^2 f_i(x). \tag{1.30}$$

The first equation is a consequence of the fact that quarks are spin-1/2 particles. In the second equation, the sum is over all quark constituents, Q_i is the quark charge, and $f_i(x)$ is the probability of quark i having a momentum x. This probability distribution is known as the parton distribution function.

Here the simple picture of a proton as consisting of only three valence quarks breaks down. Being in a bound state, the quarks inside the proton constantly emit and receive gluons. Hence part of the momentum of the proton is not carried by the quarks. Experimental data suggests that about 54% of the proton momentum is carried by the three valence quarks. The rest is attributed to the gluons and the so-called sea quarks, virtual quarks created by gluon splitting.

This picture of the proton has consequences for *pp* scattering. Depending on the center of mass energy of the interaction, the largest contribution of the *pp* collision cross section can come from qq, $q\bar{q}$, qg or gg scattering. The CTEQ[11] collaboration has provided parton density functions to use as input to the simulations of *pp* collisions used in this thesis.

Underlying event

The structure of the proton and the principle of confinement has a direct consequence when studying *pp* collisions. When two protons collide, the chance of a high momentum transfer (more then a few GeV) interaction is small. The possibility of multiple hard interactions between two protons is small enough to be ignored. However, after two partons have collided, and the collision products have escaped the proton, what remains is no longer a color singlet state. Due to confinement of QCD, the remaining partons will undergo a series of soft interactions with each other and with the products of the hard interaction, to hadronize into color singlet states (mostly pions). The products of these soft interaction also end up in the detector, and this is called the Underlying Event.

Pileup

When colliding proton beams of high enough intensity, as the LHC is designed to do, the probability of two pairs of protons colliding during one bunch crossing (see chapter 2) becomes non-negligible, which means that the products of more than one collision can end-up in the detector at the same time. This effect is called pileup. Care must be taken that the analysis is not sensitive to pileup, in order to be applicable to the high intensity environment at the LHC.

1.1.6 Standard Model background to supersymmetry

Out of all the particles and interactions that are described by the SM, two particles are the main contributors to large, irreducible backgrounds to supersymmetry. This analysis will take advantage of two properties of supersymmetric models that may help distinguish it from the SM. The first is the relatively high mass of supersymmetric particles, the second the possibility of producing a lepton in their decay. These properties will be discussed in the next section. These assumptions are quite general. A low mass of supersymmetric particles would mean supersymmetry would already have been discovered. And the properties of Dark Matter are such that DM particles seem to couple only to the weak interaction, which means they can couple to leptons.

The two particles in the SM that share this combination of properties are the top quark pairs, and the W boson, with are produced at the LHC with cross sections of respectively 159 pb and 16 nb at 7 TeV. Each produces at least one lepton and one neutrino in its decay, and each has a high enough mass to show up in a SUSY analysis. These two particles will be briefly discussed here.

The top quark

The top quark is unique in the quark section of the SM. First of all, its mass (172.9 ± 1.1 GeV [12]) is of a completely different order of magnitude from the other quarks (6 MeV- 5 GeV). This high mass makes it so unstable that it is the only quark that decays freely, i.e. before hadronization, as there is no time for hadronization to take place. Top quark decay is shown in Figure 1.4. The top decays to a W-boson and a b-quark with a branching ratio of $\sim 100\%$. The W can subsequently decay to quarks (hadronic decay), or to leptons (leptonic decay). Pair produced top quarks form a major background to supersymmetry searches such as presented in this thesis. It is one of the hardest backgrounds to suppress, and a large part of this thesis will discuss methods to understand and control this background.

Single top production, where a top quark is produced together with a b-quark or a W boson also produces a signature that is comparable to the one that is looked for in the analyses of this thesis. However, the production cross section of this process is too low to be a significant background.

W with associated jet production

As mentioned in section 1.1.5, additional gluons may be radiated by any quark in the interaction. Initial and final state radiation (ISR and FSR) will play an important role at the LHC. They can lead to large numbers of extra jets in an event, This is of specific interest to this study when looking at one of the main backgrounds to SUSY, W-boson production with associated jets, see Figure 1.5 for a Feynman diagram of this process at leading order. If the W decays leptonically ($W \to e\nu$), and ISR produces a certain number of jets (cross sections for events with up to 4 or 5 jets are high enough to be significant), it becomes harder to distinguish this background from SUSY decaying to jets and leptons, which is one of the main reasons why the method described in chapter 6 was developed.

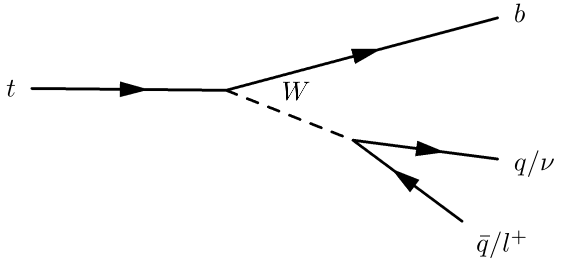


Figure 1.4: Feynman diagram of top quark decay, where the W boson may decay leptonically or hadronically.

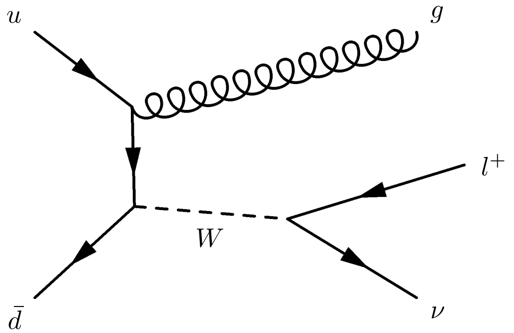
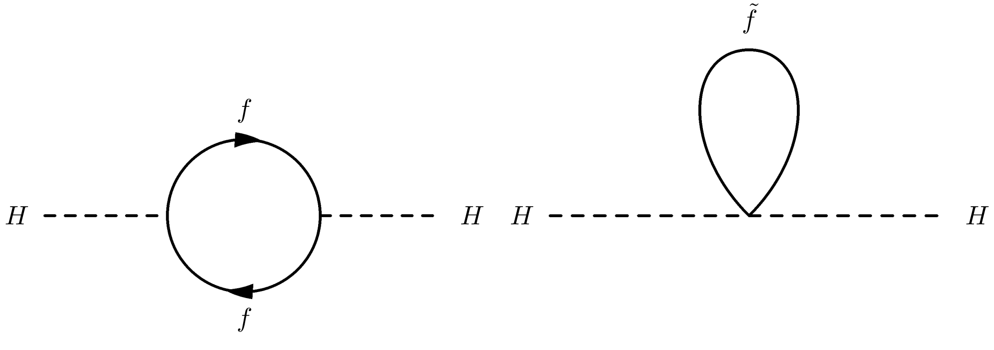


Figure 1.5: One possible Feynman diagram of W production, where the W boson decays leptonically. The gluon radiated from the quark in the initial state may produce an extra associated jet in the event.

1.2 Supersymmetric extension of the Standard Model

The Standard Model describes in great detail and with remarkable precision most, if not all, experimental results from high energy and high precision tests so far. There are however considerations both from theory and experiment that prompt a search for possible extensions to the SM. The theory of supersymmetry (SUSY) is such an extension, introducing a symmetry operation that transforms fermions into bosons and vice versa. Using this symmetry it is possible to deal with many problems of the SM, both theoretical and experimental, within the framework of quantum field theory (QFT). Furthermore SUSY provides signatures that can be looked for in the pp collisions at the LHC, which is precisely the subject of this thesis. So far no evidence of supersymmetry has been found in any experiment. The consequences of supersymmetry however are too intriguing not to look for it. Using supersymmetry as a guide, we can start exploring the possibilities of physics beyond the Standard Model, and prepare for when the first hints of new physics show up in experimental signatures.

This section will briefly discuss the key concepts of SUSY and the Minimal Supersymmetric Standard Model (MSSM). For an excellent introduction into SUSY models see the articles by Martin [13] and Aitchison [14]



(a) Fermion loop contribution to mass of the Higgs boson.

(b) Sfermion loop contribution to mass of the Higgs boson.

Figure 1.6: Contributions to the mass of the Higgs boson.

1.2.1 Common motivations for a Supersymmetric extension of the Standard Model

As said in the previous section, supersymmetry is an extension of the SM which provides a more or less elegant way of fixing some persistent theoretical issues of the SM, and at the same time makes predictions that may fit some outstanding experimental questions. This section discusses some of these problems of the SM, and how SUSY may solve them.

Hierarchy problem

The biggest theoretical problem of the SM that is tackled by SUSY is called the hierarchy problem. From fits to available electroweak data [15], the mass of the Higgs boson is $114.4 < m_H < 154.9\,\mathrm{GeV}$ evaluated at the 2σ interval, see Figure 1.7. Yet no matter what the 'bare' value of the mass of the Higgs boson might be, there is no mechanism in the SM to prevent it from blowing up to arbitrarily high scale when quantum corrections to the bare mass are included. For the fermions and gauge bosons, chiral and gauge symmetry 'protect' the masses of these particles, since in an unbroken symmetry these masses should be zero. Breaking these symmetries (assuming the breaking is relatively soft), allows for low masses of these particles. No such protecting symmetry exists for the mass of the Higgs boson. That means that diagrams like Figure 1.6 (a) give large contributions to its value. If these contributions are not canceled by contributions at higher order, the bare mass of the Higgs boson must be carefully chosen to some other value. This fine tuning of the bare mass of the Higgs boson is deemed esthetically unfavored.

SUSY provides a solution. Since any fermion appearing in loops in the Higgs mass diagrams has a supersymmetric partner of equal mass (in unbroken symmetry), every diagram like Figure 1.6 (a) has a counterpart like Figure 1.6 (b), which has a relative minus-sign in the loop contribution. Thus the Higgs mass quantum effects are canceled order by order. In unbroken SUSY, this cancelation is exact. However since we know that SUSY must be broken (no supersymmetric partners to the SM particles have been found, so they must be heavy), this cancelation is approximate. In order for the solution of the hierarchy problem to work, the mass scale of SUSY particles must be not much larger then a few TeV. This is the scale

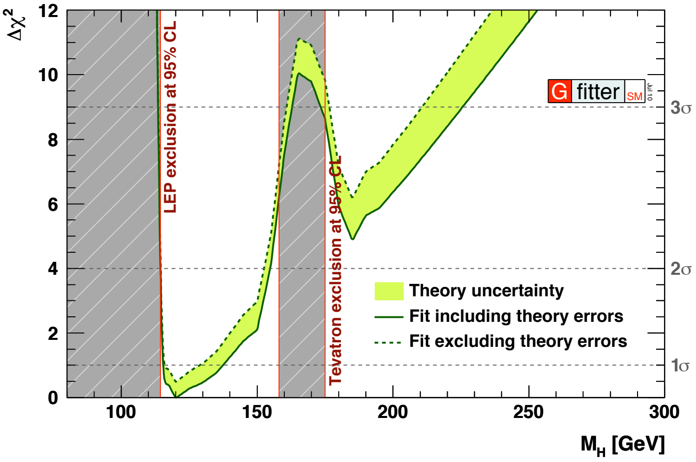


Figure 1.7: The result of the fitting procedure to all electroweak data, including the result of LEP [16] and Tevatron [17] direct Higgs searches (shaded areas). The plot shows the $(\chi^2 - \chi^2_{min})$ of the fit of all electroweak data assuming a mass for the Higgs boson plotted on the horizontal axis. Figure taken from the Gfitter collaboration [18].

that will be probed by the LHC.

Evolution of the mass of the Higgs boson (EW symmetry breaking)

The second theoretical issue that could be solved by SUSY is also related to the mass of the Higgs boson. As discussed in section 1.1.3, EW symmetry can be broken if the Higgs mass parameter μ^2 is smaller then zero. The SM provides no reason why this would be so. In the MSSM a Higgs mass defined at some high scale could decrease to a value below zero due to the running of one of the Higgs mass parameters .[1] This could provide a natural way to get a negative mass parameter at the EW breaking scale needed to obtain a v.e.v. for the Higgs field.

Grand Unification

As mentioned before, the coupling strength of the strong interaction, and in fact of all gauge interactions, depends on the energy at which a collision takes place. This running of coupling strengths has a remarkable feature: at large energy scales ($\pm 10^{16}$ GeV) the three coupling coefficients of the SM almost meet. Since they do not meet exactly, unification of the interactions strength would occur in two steps, or not at all. In the MSSM however, assuming the mass scale of the SUSY particles is around 1 TeV, the coupling strengths can unify within theoretical errors, see Figure 1.8. Historically, this has been a strong incentive for many to further investigate SUSY models.

---

[1] There is more than one Higgs doublet, and thus more than one Higgs mass parameter in the MSSM. This will be discussed later.

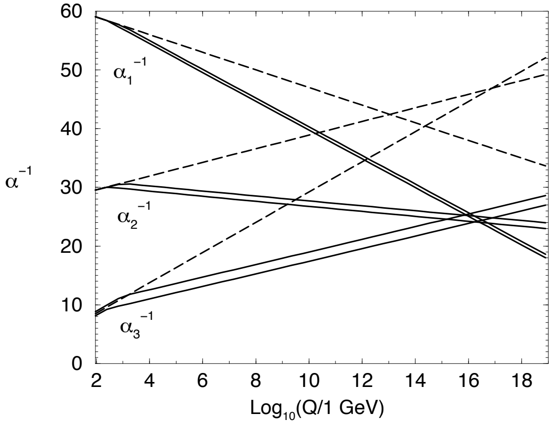


Figure 1.8: The running of coupling strengths for the three gauge interactions. The $U(1)$ coupling is denoted as α_1, $\alpha 2$ is the $SU(2)$ coupling and α_3 the strong $SU(3)$ coupling strength. The evolution is shown for the SM (dashed line) and for the MSSM (solid lines). Figure taken from Martin [13].

String theory (gravity)

Experimentally there are even fewer (if any) clues that hint towards string theory than there are for SUSY. It is however one of the few theories that promise to unite gravity with the forces of the SM. Most (and as some theorists say, phenomenologically the most interesting) string theories contain some form of low energy supersymmetry.

Dark Matter candidate

There are now many cosmological indications that there is a large amount of Dark Matter in our universe. See for instance the overview article [19]. No explanation has presented itself within the current theoretical framework for the presence of this Dark Matter. SUSY however may contain a candidate Dark Matter particle.

There exists within the framework of supersymmetry a symmetry called R-parity, a multiplicative quantum number defined as

$$P_R = -1^{3(B-L)+2s} \tag{1.31}$$

with B and L the baryon and lepton number, and s the spin quantum number. As can be easily checked this number is 1 for all SM particles and -1 for all supersymmetric partners of these particles. If this quantum number is conserved, the lightest supersymmetric particle (LSP) is stable. A large violation of R-parity would lead to rapid proton decay, making this a reasonable assumption.

In many models this LSP is the neutralino, a mass eigenstate which is a linear combination of the photon, Z and Higgs supersymmetric partners. This neutralino couples only through weak interactions, and with a mass in the range 10 GeV to about 1 TeV, it could be a good candidate for a Dark Matter particle.

Other experimental open questions

Other experimental concerns are the dominance of matter over antimatter in the universe coupled to the extremely low values of CP and baryon and lepton number violation observed at low energies. SUSY models can easily incorporate extra sources of violation of these quantum numbers, which is needed to explain baryogenesis in the early universe.

Finally, the muon anomalous magnetic moment as measured by the Brookhaven $(g-2)$ experiment [20] deviates by 2.7 σ from the expected SM value. This possible deviation might hint at new physics, and can also be explained assuming loop contributions from light SUSY [21]

1.2.2 Building a supersymmetric model

Supersymmetry transforms fermions to bosons and vice versa. The first step towards building a supersymmetric theory is to find an operator that performs this action:

$$Q|J>=|J\pm\frac{1}{2}> \tag{1.32}$$

Considering the algebra that incorporates such operators, note that since the fermionic operator Q is a symmetry operator, it commutes with the Hamiltonian, and so does the anticommutator of Q:

$$[\{Q_a,Q_b\},H]=0, \tag{1.33}$$

Thus the anticommutator is a conserved quantity. As a and b have two possible values, the anticommutator $\{Q_a,Q_b\}=Q_aQ_b+Q_bQ_a$ has three independent components, one can assume it must transform like a spin-1 object, in other words it transforms like a four-vector. There is only one conserved vector available in relativistic field theory, therefore it can be deduced from quite general arguments that the algebra should look something like:

$$\{Q_a,Q_b\}\sim P_\mu \tag{1.34}$$

Aitchison suggests [14] intriguingly that this indicates that the supersymmetry is like the square root of the momentum operator, which is represented by the space-time derivative. Thus SUSY goes "one better than the Dirac equation", by taking its square-root. Since the derivative is "locked in our notions of a four-dimensional space-time", supersymmetry effectively extends our concept of space-time as containing an extra fermionic degree of freedom. I agree with him that this is "arguably a more striking concept" than the fact that supersymmetry merely creates a superpartner for all known particles.

The exact form of the supersymmetry algebra is:

$$\begin{gathered}\{Q_a,Q_b^\dagger\}=\boldsymbol{\sigma}_{ab}P_\mu\\ \{Q_a,Q_b\}=0\quad\{Q_a^\dagger,Q_b^\dagger\}=0\end{gathered} \tag{1.35}$$

with $\boldsymbol{\sigma}_{ab}$ the Pauli matrices and the † denoting hermitian conjugation. Supersymmetric theories are generally built using spinors in the Weyl representation, i.e. the eigenstates of the chirality operator. Weyl spinors have two complex components, unlike Dirac spinors, which have four. A Dirac spinor can be written as

$$\Psi=\begin{pmatrix}\psi\\ \chi\end{pmatrix} \tag{1.36}$$

with ψ and χ the right- and left-handed spinors in the Weyl representation. It makes sense to write down a theory in terms of these spinors because of the distinction the weak force makes between left- and right-handed chirality.

A left-handed spinor can be transformed under supersymmetry by:

$$\delta_\xi \chi = -i\sigma^\mu (i\sigma_2)\xi^* \partial_\mu \phi \tag{1.37}$$

where ξ is a fermionic infinitesimal parameter, and $\sigma^\mu = (\mathbf{1}, \boldsymbol{\sigma})$, with $\boldsymbol{\sigma}$ the Pauli matrices. The reverse transformation is:

$$\delta_\xi \phi = \xi^T (-i\sigma_2)\chi \tag{1.38}$$

Using these transformations, the simplest supersymmetric Lagrangian, containing one free fermion and one free scalar

$$\mathcal{L}_{\text{chiral}} = -\partial_\mu \phi^* \partial^\mu \phi + \chi^\dagger i\bar{\sigma}^\mu \partial_\mu \chi \tag{1.39}$$

can be shown to be invariant under supersymmetry. Here $\bar{\sigma}^\mu = (\mathbf{1}, -\boldsymbol{\sigma})$. The fields ϕ and χ together form a so-called chiral supermultiplet.

There is a complication here that needs to be addressed before proceeding. Although the transformations 1.37 and 1.38 are valid as long as all particles are on-shell, off-shell the number of degrees of freedom on either side of the equation do not match. A Weyl fermion has two complex components, i.e. four degrees of freedom, while a complex scalar has only two. The on-shell equation of motion

$$\bar{\sigma}^\mu \partial_\mu \chi = 0 \tag{1.40}$$

removes two of the fermion's degrees of freedom. What is needed is an extra field that disappears on-shell (i.e. it can never be observed as a physical particle), but picks up two degrees of freedom off-shell. These auxiliary fields play in important role not only in building the supersymmetric model, but also in breaking it. For a chiral supermultiplet the auxiliary field transforms like this:

$$\delta F = -i\xi^\dagger \bar{\sigma}^\mu \partial^\mu \chi \quad \delta F^* = \partial_\mu \chi^\dagger \bar{\sigma}^\mu \xi \tag{1.41}$$

and the auxiliary field Lagrangian

$$\mathcal{L}_{\text{aux}} = F^* F \tag{1.42}$$

is added to the free field Lagrangian 1.39. The equation of motion that can be derived from this Lagrangian for the auxiliary field is $F = F^* = 0$. Thus the field vanishes on-shell, as needed. The auxiliary field also changes the transformation 1.37:

$$\delta_\xi \chi = -i\sigma^\mu (i\sigma_2)\xi^* \partial_\mu \phi + \xi F \tag{1.43}$$

Only chiral multiplets allow the left- and right-handed components of a fermion field to be treated differently like in the SM. Hence all fermions in the SM must be part of a chiral multiplet.

Gauge vector bosons are the bosonic part of a gauge supermultiplet. The superpartner of the vector boson A^a_μ is a fermion represented by a two-component L-type Weyl spinor λ^a with index a running over the adjoint representation of the gauge group. The supersymmetric transformations for these fields looks like:

$$\begin{aligned} &\delta_g A^a_\mu = \partial_\mu \Lambda^a + g f^{abc} A^b_\mu \Lambda^c \\ &\delta_g \lambda^a = g f^{abc} \lambda^b \Lambda^c \end{aligned} \tag{1.44}$$

for a general, non-abelian gauge group. In this formula, Λ^a is an infinitesimal transformation parameter, g is the gauge strength, and f^{abc} are the structure constants of the gauge group. In the case of an abelian gauge group such as U(1), $f^{abc} = 0$. The simplest Lagrangian for the components of a gauge supermultiplet can be written in terms of the Yang-Mills field strength:

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu. \tag{1.45}$$

Furthermore, as with the chiral supermultiplets, there is a mismatch in degrees of freedom of the two fields, when they are considered off-shell. The fermionic field λ^a has two complex components and four degrees of freedom. The bosonic field A^a_μ has only three, since one is removed by gauge invariance. Therefore a real bosonic auxiliary field called D^a is coupled to the supermultiplet, resulting in the non-interacting gauge part of the Lagrangian :

$$\mathcal{L}_{\text{gauge}} = \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu a} + i\lambda^{\dagger a} \bar{\sigma}^\mu D_\mu \lambda^a + \frac{1}{2} D^a D^a. \tag{1.46}$$

Together equations 1.39 (with ∂_μ replaced by a covariant derivative), 1.42 and 1.46 give the general form of a supersymmetric Lagrangian, except for the Yukawa couplings between scalars and fermions. It includes all SM fermions and Higgs particle as one part of a chiral supermultiplet and all the SM gauge bosons combined with a fermionic superpartner in gauge supermultiplets. The complete particle content in the MSSM are discussed in the next sections.

When more than one chiral multiplet exists within a theory, non-gauge (Yukawa) interactions are introduced in supersymmetric theories through:

$$\mathcal{L}_{\text{int}} = \left(-\frac{1}{2} W^{ij} \chi_i \chi_j + W^i F_i \right) + h.c. \tag{1.47}$$

with indices i and j running over the available multiplets, and where the quantities W^{ij} and W^i derive from a polynomial in the scalar fields ϕ_i called the superpotential:

$$\begin{aligned} W^{ij} &= \tfrac{\delta^2}{\delta\phi_i \delta\phi_j} W \\ W^i &= \tfrac{\delta}{\delta\phi_i} W \end{aligned} \tag{1.48}$$

The exact form of the superpotential depends on the theory. A general form that is applicable to the MSSM consists of a fermion mass term and a Yukawa coupling:

$$W = \frac{1}{2} M_{ij} \phi_i \phi_j + \frac{1}{6} y^{ijk} \phi_i \phi_j \phi_k \tag{1.49}$$

Using the equations of motion for the auxiliary field when interactions are included ($F_i = -W^*_i$ and $F^{*i} = -W^i$), we can write down the scalar potential of the supersymmetric Lagrangian:

$$\begin{aligned} &V(\phi, \phi^*) = W^k W^*_k = F^{*k} F_k = \\ &M^*_{ik} M^{kj} \phi^{*i} \phi_j + \tfrac{1}{2} M^{in} y^*_{jkn} \phi_i \phi^{*j} \phi^{*k} + \tfrac{1}{2} M^*_{in} y^{jkn} \phi^{*i} \phi_j \phi_k + \tfrac{1}{4} y^{ijn} y^*_{kln} \phi_i \phi_j \phi^{*k} \phi^{*l} \end{aligned} \tag{1.50}$$

Putting it all together we can write for the Lagrangian

$$\begin{aligned} &\mathcal{L} = -\partial_\mu \phi^{*i} \partial^\mu \phi_i + \chi^{\dagger i} i \bar{\sigma}^\mu \partial_\mu \chi_i - V(\phi, \phi^*) + \\ &\tfrac{1}{2} M^{ij} \chi_i \chi_j + \tfrac{1}{2} M^*_{ij} \chi^{\dagger y} \chi^{\dagger j} + \tfrac{1}{2} y^{ijk} \phi_i \chi_j \chi_k + \tfrac{1}{2} y^*_{ijk} \phi^{*i} \chi^{\dagger j} \chi^{\dagger k} \end{aligned} \tag{1.51}$$

This simplest form of a supersymmetric Lagrangian with only chiral supermultiplets shows

- Yukawa interaction between scalars and fermions (y^{ijk} terms)
- scalar self-interaction (ϕ^4 terms)
- scalar and fermion mass terms (M^{ij} terms)

where it should be noted that the mass terms for the scalars and the fermions have the same coefficient, which shows that in unbroken supersymmetry the particles within a supermultiplet must be of equal mass.

Gauge interactions are introduced in the supersymmetric Lagrangian by replacing all derivatives with the covariant derivative as usual. However, the other two components of the gauge supermultiplet (the gaugino λ^a and the auxiliary field D^a) obviously also couple to the chiral fields. This is how [2] :

$$\mathcal{L}_{\text{gaugino}} = -\sqrt{2}g(\phi^* T^a \chi)\lambda^a - \sqrt{2}g\lambda^{\dagger a}(\chi^\dagger T^a \phi) + g(\phi^* T^a \phi)D^a \tag{1.53}$$

Here T^a are the generators of the gauge group, and g is the gauge coupling strength. These Lagrangians define all (self-) couplings and mass terms that can be found in the MSSM. The precise form of the couplings depend of course on the particle content and the form of the superpotential. These will be discussed in the next section.

1.2.3 Particle content of the MSSM

Looking at the particles discovered so far, it is clear that if supersymmetry is realized in nature, it must be broken. From the previous section we know that for unbroken superpartners must be of equal mass and equal gauge group representation, and no such partners with different spin have been discovered. At the same time it is clear that none of the particles discovered so far can be each others superpartner, as their respective gauge quantum numbers do not match. The only possible match would be of the Higgs scalar with one of the neutrinos, but as it turns out this would lead to large lepton number violation which is not realized in nature.

Hence the MSSM introduces a new supersymmetric particle (sparticle) for each of the known SM particles. The naming convention of superpartners to SM particles is such that the partners of a fermion get an 's-' in front of their name, so the partners of electron, tau and quarks are the selectron, stau and squarks respectively. The partners of the bosons get a '-ino' suffix, giving the gluino and the Wino for the partners of the gluon and the W. A sfermion thus is a bosonic partner to a fermion, and a gaugino a fermionic partner to a gauge boson.

The as yet undiscovered Higgs boson is a bit more complex. Anomaly cancelation requires that there are two Higgs $SU(2)$ multiplets in the MSSM, with a superpartner called a Higgsino for each of the four fields. Two of the four Higgs fields have an electric charge. When $SU(2) \times U(1)$ is broken through the Higgs mechanism, the 'up-type' Higgs field H_u and the down-type Higgs H_d obtain a v.e.v. v_u and v_d respectively. The coupling of the H_u to the up-type quarks, H_d to the down-type quarks and electrons gives these particles their masses. One of the vevs is fixed by the experimental values of the W and Z mass, leaving only the ratio v_u/v_d as a free parameter of the theory. This ratio is often quoted as the tangent of an angle, $\tan\beta$. The

---

[2] Apart from replacing derivatives with covariant derivatives and this addition of gaugino and auxiliary coupling, the transformation of the chiral auxiliary field is changed:

$$\delta F_i = -i\xi^\dagger \bar{\sigma}^\mu \partial^\mu \chi + \sqrt{2}g(T^a\phi)_i \xi^\dagger \lambda^{\dagger a} \tag{1.52}$$

Name	Symbol	particle	sparticle	representation
quarks, squarks	Q	(u_L, d_L)	$(\tilde{u}_L, \tilde{d}_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	u	u_R	$\tilde{u}_R$	$(\mathbf{3}, \mathbf{1}, -\frac{2}{3})$
	d	d_R	$\tilde{d}_R$	$(\mathbf{3}, \mathbf{1}, \frac{1}{3})$
leptons, sleptons	L	(ν, e_L)	$(\tilde{\nu}, \tilde{e}_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	e	e_R	$\tilde{e}_R$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, Higgsinos	H_u	(H_u^+, H_u^0)	$(\tilde{H}_u^+, \tilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, \frac{1}{2})$
	H_d	(H_d^0, H_d^-)	$(\tilde{H}_d^0, \tilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
gluon, gluino		g	$\tilde{g}$	$(\mathbf{8}, \mathbf{1}, 0)$
W, wino		$W^\pm, W^0$	$\tilde{W}^\pm, \tilde{W}^0$	$(\mathbf{1}, \mathbf{3}, 0)$
B, bino		B^0	$\tilde{B}^0$	$(\mathbf{1}, \mathbf{1}, 0)$

Table 1.1: Table of all SM particles with their supersymmetric counterparts. The $SU(3)_C \times SU(2)_L \times U(1)_Y$ representation to which all fields in one supermultiplet belong is also listed. The B and bino of course refer to the unbroken EW bosons, the first symbol after the fermion and Higgs names refer to the notation of the fields in the MSSM Lagrangian.

Higgs fields obtain a v.e.v. after supersymmetry is broken, so $\tan\beta$ is not a parameter of the unbroken MSSM.

All SM particles with their superpartners are listed in Table 1.1. With the field content of the MSSM of Table 1.1 only the superpotential is needed to complete the MSSM theory:

$$W = \bar{u}\mathbf{y_u}QH_u - \bar{d}\mathbf{y_d}QH_d - \bar{e}\mathbf{y_e}LH_d + \mu H_u H_d \tag{1.54}$$

where the symbols denote the superfields of Table 1.1. The coupling parameters $\mathbf{y}$ are the same Yukawa couplings as in the SM. This superpotential introduces more couplings to the MSSM that are not the SM gauge couplings or their supersymmetric counterparts. With respect to the supersymmetric counterparts of the SM gauge couplings, these are easy to incorporate (assuming R-parity is conserved): the MSSM gaugino/gauge/sfermion/fermion couplings can be obtained from the SM gauge coupling vertices replacing two or four outgoing lines with their supersymmetric partner.

With the particle content of Table 1.1 and this superpotential, the unbroken MSSM is now fully specified. Note that the MSSM has introduced only one new parameter, μ.

1.2.4 Supersymmetry breaking

As no SUSY partner candidates to any SM particles have been discovered, supersymmetry must be broken. As the symmetry is the basis of the theory under construction, it seems natural to look for a method of spontaneous SUSY breaking, meaning that the Lagrangian remains invariant under SUSY. However, there is no hint or clue as to how the spontaneous breaking of SUSY occurs. There are plenty of ideas, but so far there is no way of choosing one over the other.

Supersymmetry breaking terms

So instead of implementing a formal SUSY breaking mechanism, the result of some SUSY breaking at large scale is inserted into the Lagrangian, breaking SUSY explicitly, reducing the MSSM to an effective, low energy theory. These SUSY violating terms should be 'soft', i.e. have coupling dimensions of mass or higher. Non-soft terms would re-introduce quadratic divergences to the mass of the Higgs boson, thus negating the solution to the hierarchy problem that SUSY was designed to fix. The Lagrangian with soft SUSY breaking terms looks like:

$$\begin{aligned}\mathcal{L}_{\text{soft}} \quad = \quad & -\tfrac{1}{2}(M_3\tilde{g}\tilde{g} + M_2\tilde{W}\tilde{W} + M_1\tilde{B}\tilde{B} + c.c.) \\ & -(\tilde{\bar{u}}\mathbf{a_u}\tilde{Q}H_u - \tilde{\bar{d}}\mathbf{a_d}\tilde{Q}H_d - \tilde{\bar{e}}\mathbf{a_e}\tilde{L}H_d + c.c.) \\ & -\tilde{Q}^\dagger\mathbf{m^2_Q}\tilde{Q} - \tilde{L}^\dagger\mathbf{m^2_L}\tilde{L} - \tilde{\bar{u}}\mathbf{m}^2_{\bar{\mathbf{u}}}\tilde{\bar{u}}^\dagger - \tilde{\bar{d}}\mathbf{m}^2_{\bar{\mathbf{d}}}\tilde{\bar{d}}^\dagger - \tilde{\bar{e}}\mathbf{m}^2_{\bar{\mathbf{e}}}\tilde{\bar{e}}^\dagger \\ & -m^2_{H_u}H^*_uH_u - m^2_{H_d}H^*_dH_d - (bH_uH_d + c.c.)\end{aligned} \tag{1.55}$$

The last line of this equation shows separate mass parameters for the H_u and H_d supermultiplets. These will play an important role in EW symmetry breaking in the MSSM.

Supersymmetry universality

The unbroken MSSM has not added a great deal of new parameters with respect to the SM. Apart from the μ-term, all couplings and masses are fully determined by SM parameters. This nice feature has been removed from the model with SUSY breaking, which brings the total to the 105 extra parameters of the model. This seems a bit much, and one might state that with so much new parameters, the new model is completely arbitrary. However, most of the new parameters are already severely limited by experimental measurements of CP-violation or flavor mixing. An arbitrary MSSM model therefore cannot work. There is a nice simplification which circumvents the CP-violation and flavor changing effects of the MSSM, which is the principle of SUSY universality. This means that the slepton and squark mass matrices are diagonal in family space, to avoid lepton flavor violation, that the tri-linear couplings are proportional by a real number to the Yukawa couplings of the unbroken MSSM. And finally also the gaugino masses must be approximately real, thus no new CP-violating phase is introduced.

It is of course legitimate to impose these demands on a supersymmetric theory from experimental constrains. However the theory would be arguably more convincing if some model of spontaneous SUSY breaking explains the origin of SUSY universality. Several such models exist. I will shortly name a few, and afterwards look in more detail at the specific model that was used in this analysis.

1.2.5 Models of high-energy supersymmetry breaking

Spontaneous symmetry breaking implies that the vacuum is no longer invariant under the symmetry. For the supersymmetry operator Q this means:

$$Q_a|0> \neq 0, \qquad Q^\dagger_b|0> \neq 0. \tag{1.56}$$

This has an interesting consequence. As the anticommutator of supersymmetry is related to four-momentum as in equation 1.35, the Hamiltonian (with expectation value p_0) can be expressed in terms of these anticommutators of supersymmetry operators. It can then be shown that spontaneous supersymmetry breaking is equivalent to a non-zero energy term of the vacuum. See for instance section 15.1 of the article by Aitchison, [14]. Assuming the vacuum is a scalar, this means that the vacuum expectation value of the scalar potential is non-zero:

$$< 0|V|0 >=< 0|(F^{*i}F_i + \frac{1}{2}\sum_a D_aD_a)|0 >, \tag{1.57}$$

with equation 1.50 inserted for the scalar potential. This means that in order to break supersymmetry, one of the auxiliary fields must obtain a v.e.v.. D-term breaking works only if the auxiliary field is part of a $U(1)$ supermultiplet. Since in the process of SUSY breaking the corresponding gauge boson would require a mass, this multiplet cannot be that of electromagnetism. Even when a new $U(1)$ gauge field is created, it seems difficult to produce a workable phenomenology.

The SUSY breaking model under investigation in this thesis is therefore based on creating a v.e.v. for some F-term. This must be the auxiliary field to a gauge-singlet chiral supermultiplet, in order for the vacuum state to be uncharged. Furthermore it is experimentally excluded that the breaking can be created using only renormalizable interactions at tree level: the masses of the particles that can be generated are so light that they would have been discovered already. Therefore, F-type supersymmetry breaking requires a new sector, where the auxiliary field of a gauge singlet in that sector communicates to the MSSM particles only indirectly through loop interactions.

There are two main scenarios that accomplish this. The first, gauge mediated supersymmetry breaking, assumes that the MSSM breaking terms arise from interactions with messenger particles through the usual $SU(3) \times SU(2) \times U(1)$ interaction. These would couple via some new interaction to the gauge singlet F-term, thus providing their role as messenger.

The other model is called Planck scale mediated supersymmetry, or supergravity. Gravity is introduced into the MSSM by making SUSY a local symmetry: the infinitesimal parameter ξ of the SUSY transformations is now a function of the space-time coördinates $\mathbf{x}$. The quantum particle of gravity gets a supersymmetric partner, the gravitino, which acts in a similar way as the gauge fields of the gauge interactions. Through a mechanism called super-Higgs the gravitino becomes massive when SUSY is broken, completely analogous the the Higgs mechanism of the EW theory. This thesis uses simulations of this type of supersymmetry models, so it will be discussed further.

When the hidden sector where one of the chiral multiplets obtain an auxiliary field with a non-zero v.e.v. communicates through gravity with the particles of the MSSM there will be interaction terms in the MSSM Lagrangian that look like:

$$\mathcal{L}_{SUGRA} = -\frac{1}{M_P}F(\frac{1}{2}f_a\lambda^a\lambda^a + \frac{1}{6}y'^{ijk}\phi_i\phi_j\phi_k + \frac{1}{2}\mu'^{ij}\phi_i\phi_j) + c.c. - \frac{1}{M_P^2}FF^*k^i_j\phi_k\phi^{*j} \quad (1.58)$$

where the terms are of the same form as equation 1.55. Here the interaction has dimension 1/[energy] and is determined by the Planck scale M_P at which gravity becomes important. The minimal form of this type of supersymmetry breaking (minimal supergravity or mSUGRA) assumes a limited number of values can determine these parameters at the GUT scale. This means that in formula 1.58 $f_a = f$, $k^i_j = k\delta^i_j$, $y'^{ijk} = \alpha y^{ijk}$ and $\mu'^{ij} = \beta\mu^{ij}$. Rewriting the terms of equation 1.58:

$$m_{1/2} = f\tfrac{<F>}{M_p} \quad m_0^2 = k\tfrac{|<F>|^2}{M_p^2} \quad A_0 = \alpha\tfrac{<F>}{M_P} \quad B_0 = \beta\tfrac{<F>}{M_P} \quad (1.59)$$

This means a huge simplification of the terms in equation 1.55. At the GUT scale, all gauginos and Higgs' get a mass $m_{1/2}$, all sleptons and Higgsinos a mass of m_0, and the b-term is equal to $B_0\mu$. The trilinear couplings are all proportional to the MSSM Yukawa couplings by a factor A_0. Together with the μ term from the unbroken MSSM, this leaves mSUGRA to be determined by only 5 parameters. Incorporating EW symmetry breaking will fix the absolute value of μ, reducing the total number of mSUGRA to 4 plus a sign. This gives a workable starting point to start exploring the MSSM phase space. To do so, I will discuss the phenomenology of the mSUGRA framework in the next section.

1.2.6 Phenomenology of the mSUGRA framework

In this section I will discuss the new particles that are visible according to the MSSM: what particles can be created and measured at the LHC and the ATLAS detector. I will discuss the Higgs sector, the gauginos and the sleptons and squarks. Starting from the high energy model such as discussed in the previous section, somehow the parameters that define the specific form of the mSUGRA model under study have to be transported down to the energy at which we look for them. This means going from the GUT scale down to order 1 TeV in energy. In QFT, the evolution of a parameter is calculated using the renormalization group equations (RGE). Both Martin [13] and Aitchison [14] provide excellent discussion on how to derive these equations, as well as the exact form at one loop order. In this section I will only quote the qualitative results of these discussions. What is important to realize is that although the theory at high energy is defined by only four parameters and a sign, the distinct properties of individual supermultiplets and their component fields result in a diverse phenomenology at low energies.

As stated in section 1.2.1, there is a mechanism in the MSSM that provides a natural way for the Higgs potential to obtain a v.e.v., thus breaking EW symmetry. As the MSSM contains two Higgs multiplets, the details of the Higgs mechanism are a bit more involved then for the SM. There are however also some similarities. Of the available degrees of freedom of the Higgs (now a total of eight), three are still used to provide mass to the Z and $W^{\pm}$. It is possible, and since electromagnetism is unbroken also preferable, to set the v.e.v. of the charged Higgs' to zero. If only the neutral H_u^0 and H_d^0 obtain a v.e.v. (v_u and v_d, respectively), they can both be made real, and positive. In order to obtain a non-zero v.e.v. for both fields, either $m_{H^2_d}$ or $m_{H^2_u}$ must be negative. The latter can become negative in a natural way, since the RGE of this parameter is such that its value decreases with decreasing energy, caused by a large coupling to the (heavy) top quark, which H_d does not have. Thus it is possible that $m_{H^2_u}$ goes from a positive value at high energies to a negative value at low energies, while $m_{H^2_d}$ and all sfermion masses which all start at the same m_0 at the GUT scale remain positive. Finding the minimum of the Higgs scalar potential by setting:

$$\frac{\partial V}{\partial H_u} = \frac{\partial V}{\partial H_d} = 0 \tag{1.60}$$

gives the following equations for the Higgs v.e.v. and mass parameters:

$$\begin{aligned} (|\mu|^2 + m_{H^2_u})v_u &= bv_d + \tfrac{1}{4}(g^2 + g'^2)(v_d^2 - v_u^2) \\ (|\mu|^2 + m_{H^2_d})v_d &= bv_u - \tfrac{1}{4}(g^2 + g'^2)(v_d^2 - v_u^2) \end{aligned} \tag{1.61}$$

One can then use the masses of the Z in terms of the $U(1)$ coupling g, the $SU(2)$ coupling g' and the Higgs vevs to eliminate some parameters:

$$m_Z^2 = \frac{1}{2}(g_2 + g'^2)(v_u^2 + v_d^2) \tag{1.62}$$

which gives

$$\begin{aligned} (|\mu|^2 + m_{H^2_u}) &= b\cot\beta + \tfrac{m_Z^2}{2}\cos 2\beta \\ (|\mu|^2 + m_{H^2_d}) &= b\tan\beta - \tfrac{m_Z^2}{2}\cos 2\beta \end{aligned} \tag{1.63}$$

These equations can be used to eliminate b and μ from the list of free parameters of the MSSM in favor of $\tan\beta$. The phase of μ remains undetermined, and since this parameter is usually taken to be real in order to avoid large CP violation effects, it is the sign of μ that is a free parameter of mSUGRA.

This does not fully determine the mass of the Higgs field however. Expanding the potential around its minimum mixes different states. The imaginary parts of the neutral Higgses form one massless state ('eaten' by the Z), and a heavy state A_0 with mass $m_{A^0} = (2b/\sin 2\beta)^{\frac{1}{2}}$. The charged Higgses give rise to two massless states (providing the longitudinal d.o.f. of the $W^{\pm}$), and two charged Higgs states with mass $m_{H^{\pm}} = \sqrt{m_W^2 + m_{A^0}^2}$. The real parts of the neutral Higgses finally give rise to two states, a light Higgs h^0 and a heavy Higgs H_0 with masses

$$m_{h^0} = \frac{1}{2}\left\{m_{A^0}^2 + m_Z^2 - [(m_{A^0}^2 + m_Z^2)^2 - 4m_{A^0}^2 m_Z^2 \cos^2 2\beta]^{\frac{1}{2}}\right\} \tag{1.64}$$

$$m_{H^0} = \frac{1}{2}\left\{m_{A^0}^2 + m_Z^2 + [(m_{A^0}^2 + m_Z^2)^2 - 4m_{A^0}^2 m_Z^2 \cos^2 2\beta]^{\frac{1}{2}}\right\} \tag{1.65}$$

The most interesting aspect of the first of these equations is that it has a maximum:

$$m_{h^0} \leq m_Z |cos 2\beta| \tag{1.66}$$

at tree level. At higher levels, this upper limit is extended somewhat, but the prediction remains that $m_{h^0} \leq 140\,\mathrm{GeV}$, indicating that if no Higgs-like particle is discovered at the LHC, the MSSM can be completely excluded.

The next set of particles to discuss are the neutralinos. These are mass eigenstates formed by mixing the two neutral Higgsinos with the neutral gauge bosinos of unbroken EW theory, the neutral Wino and Bino. The neutralinos are called $\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$, $\tilde{\chi}_3^0$ and $\tilde{\chi}_4^0$ in order of ascending mass (by convention). Two effects contribute to the mixing of these states. One is the bilinear terms involving one of the neutral gauginos with one of the neutral Higgsinos after EW symmetry breaking (a consequence of the Yukawa coupling to the gauginos). The second is the μ terms for the Higgsinos mixing $\tilde{H}_u$ and $\tilde{H}_d$. It is the lightest neutralino which is a prime (although not the only) candidate for a Dark Matter particle in mSUGRA models, assuming it is the lightest supersymmetric particle (LSP).

Like their neutral counterparts, the charged Winos and charged Higgsinos also mix. There are four mass eigenstates as a result of this mixing, the $\tilde{\chi}_{1,2}^+$ coming from $\tilde{W}^+$ and $\tilde{H}_u^+$, and the $\tilde{\chi}_{1,2}^-$ from $\tilde{W}^-$ and $\tilde{H}_u^-$.

The evolution of the gauginos is sensitive to the coupling strengths in such a way that the gluinos are generally thought to be much more massive then the charginos and neutralinos. Since they are the only color-octet fermion in the MSSM, they do not mix with the other gauginos and the Higgsinos.

The last, and largest group of new particles in the MSSM are the squarks and sleptons. Here the first and second generation are generally treated separately from the third. Specifically, the Yukawa couplings of the first two generations are so small that they are neglected in the RGEs. This means that the mass evolution of these squarks and sleptons from the GUT scale down to 1 TeV is determined only by gauge interactions. Because L-type and R-type sfermions have different gauge group representation, they will be treated separately, and will in general have different masses at low energy. Each RGE contribution from gauge interactions increases the mass of the sparticle approximately proportional to the interaction strength. This means that it is likely that the squarks are heavier than the sleptons. Another contribution to their mass comes from EW symmetry breaking. This term splits the down-type sfermion mass from the up-type sfermion mass, which in most cases is the lighter of the two.

There are other contributions, proportional to the Yukawa coupling and therefore only relevant for the third family. For instance the left-handed stop-stop-Higgs term will result in a mass term for the stop after the Higgs field H_u^0 obtains a v.e.v.. This term is proportional to $v_u^2 y_t^2 = m_t^2$. The same applies to the right-handed stop.

Also unique to the third family are mixing terms between L- and R-type sfermions, which come from L/R mixing Yukawa couplings $-|y_t \tilde{t}_R^\dagger \tilde{t}_L - \mu v_d|^2$. Working out the square gives off-diagonal terms:

$$\mu v_d y_t (\tilde{t}_R^\dagger \tilde{t}_L + \tilde{t}_L^\dagger \tilde{t}_R) = \mu m_t \tan\beta (\tilde{t}_R^\dagger \tilde{t}_L + \tilde{t}_L^\dagger \tilde{t}_R) \tag{1.67}$$

with similar terms for the $\tilde{b}$ and $\tilde{\tau}$. Similar terms can come from the trilinear SUSY-breaking terms, which give :

$$-A_0 v_u y_t (\tilde{t}_R^\dagger \tilde{t}_L + \tilde{t}_L^\dagger \tilde{t}_R) = -A_0 m_t (\tilde{t}_R^\dagger \tilde{t}_L + \tilde{t}_L^\dagger \tilde{t}_R) \tag{1.68}$$

Thus instead of a $\tilde{t}_R$ and a $\tilde{t}_L$ states the mSUGRA models contain the mass eigenstates $\tilde{t}_1$ and $\tilde{t}_2$. The same holds for the stau and sbottom. For the stop this mixing is likely to be large due to the large top mass in the off-diagonal terms of the mass matrix. This results in the mass of the lighter stop to be much smaller than the mass of any other squark. For high values of $\tan\beta$, the mixing for the sbottom and stau will also be quite large, resulting in low masses for these sparticles relative to their first and second family counterpart.

1.2.7 Experimental signatures of mSUGRA

The phenomenology described above leads to very distinct features that set SUSY apart from SM backgrounds. Before detecting SUSY particles at the LHC, they must first be created in *pp* collisions. By far the largest fraction of SUSY production will be through pair-creation of strongly interacting particles. Depending on the SUSY mass spectrum, this will be $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$, $\tilde{q}\tilde{q}$ or $\bar{\tilde{q}}\tilde{q}$ production. The produced gluinos and squarks will decay until the LSP is produced (assuming R-parity conservation). A neutralino LSP, which interacts only weakly will escape the detector unseen, much like a neutrino. This results in a large average $E_{\mathrm{T}}^{\mathrm{miss}}$ for events where SUSY particles are created. This large missing energy is the cornerstone for most of the mSUGRA based SUSY searches at the ATLAS detector. They are further distinguished by the number of leptons they require in the final state.

Zero-lepton analyses focus only on the escaping dark-matter candidate, and thus are not very SUSY specific. The cross section (based on mSUGRA models) of events with large $E_{\mathrm{T}}^{\mathrm{miss}}$, large numbers of jets and no leptons is extremely large. Unfortunately, so is the SM background contribution in this channel. Furthermore the uncertainty on the QCD cross section makes this method subject to large systematic errors.

One-lepton searches such as the one described in this thesis make use of the fact that SUSY models couple to the weak interaction, which means they are likely to produce a single lepton (plus a neutrino). This suppresses a large fraction of the background (mainly QCD), as will be discussed in chapter 6. This type of search is not very SUSY specific.

A signal which is SUSY-specific is the same-sign dilepton signal. This makes use of the fact that the gluinos that are produced are Majorana particles, i.e. they are their own anti-particle. Each gluino can decay to a chargino with either sign, which then decays to a lepton and a neutralino. If both gluinos decay along this path, the second lepton has equal chance of being like-sign as opposite sign. The like-sign dilepton signal has very little SM background. At the same time, the cross sections for these type of decays in the mSUGRA framework are low.

Even with only 5 parameters the variety of signatures from mSUGRA models is large, and the range of allowed values for its parameters almost endless. In order to get a measure of the different behavior of mSUGRA in different part of its parameter space, for each analysis in this thesis a grid of points was defined in the area of phase space that the analysis was sensitive to. This allowed experimenters to get an idea of what a SUSY signal could look like, and to compare the SUSY prediction to experimental data. Figure 1.9 shows the parts of mSUGRA phase space that have been excluded by previous experiments. The exclusion at 95% C.L. is given as a function of m_0 and $m_{1/2}$, for given values of A_0, $\tan\beta$ and $sgn(\mu)$.

Finally, note that the analysis in this chapter makes as little assumptions on the details of new physics as possible. Any strongly produced new physics theory which has an invisible Dark Matter candidate with a typical mass around 100 GeV, and which can produce a lepton in their decay has the possibility to show up in this analysis. The goal of this analysis is therefore not proving the existence of supersymmetry, but rather uses supersymmetry as a starting point for the search for new physics.

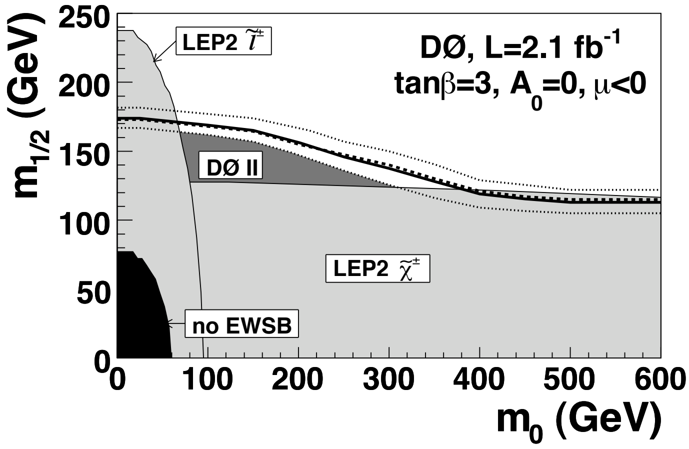


Figure 1.9: The parts of mSUGRA phase space as a function of the scalar and fermion mass parameters m_0 and $m_{1/2}$ which were excluded at 95% C.L. by previous experiments. Figure taken from [22].

Chapter 2

LHC and the ATLAS detector

2.1 The LHC

The Large Hadron Collider (LHC) is designed to collide two beams of protons, at an energy of 7 TeV per beam and a design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. A detailed description of the LHC can be found in the LHC design paper [23]. The machine is located in a circular tunnel of 27 km length approximately 100 m underground, between the airport of Geneva and the Jura mountains. Figure 2.1 shows the location of the LHC tunnel, and the locations of the four large experiments around its ring.

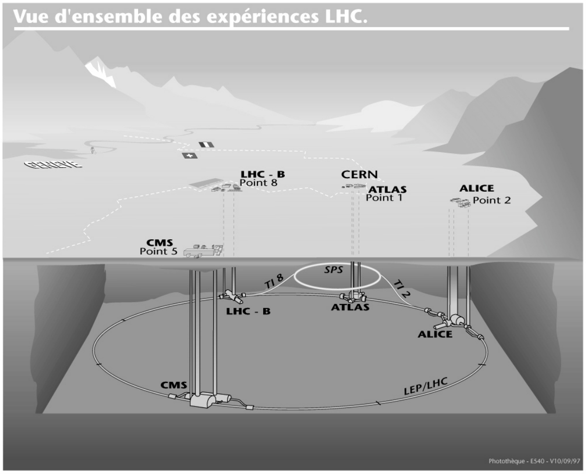


Figure 2.1: The LHC accelerator and the four largest experiments located around its ring, 100 m underground between Geneva airport and the Jura mountains.

The bending power of the LHC is provided by superconducting dipole magnets, which produce a 8.33 T magnetic field at full energy. There are four interaction points along the ring, where the beams are made to collide. At these points, the LHC experiments are located. Two of these, ATLAS[24] and CMS[25], are general purpose detectors, designed to look for expected and unexpected physics at the highest possible energy. The LHCb[26] experiment is built to precisely measure the properties of B-mesons, looking e.g. at the nature of CP violation. The LHC also has the capability of colliding heavy ions instead of protons. The ALICE[27] experiment is designed to look at the result of these collisions, investigating the new state of matter, the so-called Quark Gluon Plasma. The TOTEM[28] experiment is placed on either side of the CMS detector at a distances of 147 m and 220 m. It is designed to measure the total pp cross section and to study elastic and diffractive scattering. The LHCf[29] experiment, located 140 m away from the ATLAS interaction point on either side, is designed to measure the very forward production of neutral particles from pp collisions, which is valuable input to air shower models for cosmic ray detection.

2.2 The ATLAS detector

The ATLAS detector at point 1 on the LHC ring is one of the two general purpose experiments. A schematic view of the detector is shown in Figure 2.2. It was designed to be able to detect signs of the Higgs boson irrespective of its mass, or new physics such as supersymmetry. High resolution in the measurements of the energy and direction of particles are needed to distinguish signals from new physics from Standard Model backgrounds, such as QCD production of jets, W- and Z-boson production and $t\bar{t}$ pair production.

The inner detector provides high resolution tracking and vertex reconstruction, which helps to identify long lived particles. This is used in the $H \to b\bar{b}$ channel, and in the search for some models beyond the Standard Model. The calorimeters provide high resolution energy measurement for electrons, photons and jets and can be used for particle identification of photons, electrons, muons and taus. It also provides a near hermetic coverage for a good measurement of the total energy, as well as possible missing energy. Outside the calorimeter systems, the muon spectrometer provides triggering, identification and momentum measurement of muons. A solenoid magnet around the inner detector and a toroid magnet within the muon spectrometer provide bending power of tracks for momentum measurement.

The coordinate system of ATLAS is defined as follows: the x-axis points from the interaction point to the center of the ring, and the y-axis points upwards. The z-axis points along the beam axis. The direction of positive z points towards the A-side of ATLAS (in the direction of Geneva airport), the negative z direction points towards the C-side. The Barrel (B) part of ATLAS thus lies in between A and C side endcaps. The azimuthal angle around the beam axis is called ϕ, and the polar angle with respect to the positive z axis θ. The momentum in the plane perpendicular to the beam axis is given by $p_T = p \sin\theta$. The polar angle is usually measured in units of pseudorapidity $\eta = \ln(\tan(\theta/2))$. This variable is better suited for use in collider experiments since the density of particles from a collision is approximately equal per unit of η for all values of η and $\Delta\eta$ is invariant under boosts in de z direction.

Each subdetector is divided in two parts. The central region ($|\eta| \lesssim 1.5$) has a barrel like geometry, and is therefore called the barrel. On both sides of the barrel the detectors are arranged in one or several disks, providing the endcap to make sure that all directions around the interaction point are fully covered.

For each bunch crossing (event), all subsystems will measure the outcome of the pp collision. Since the design rate of collisions (40MHz) is too high to store all events, some fast detector systems sent the first information to the trigger system. This system consists of three stages,

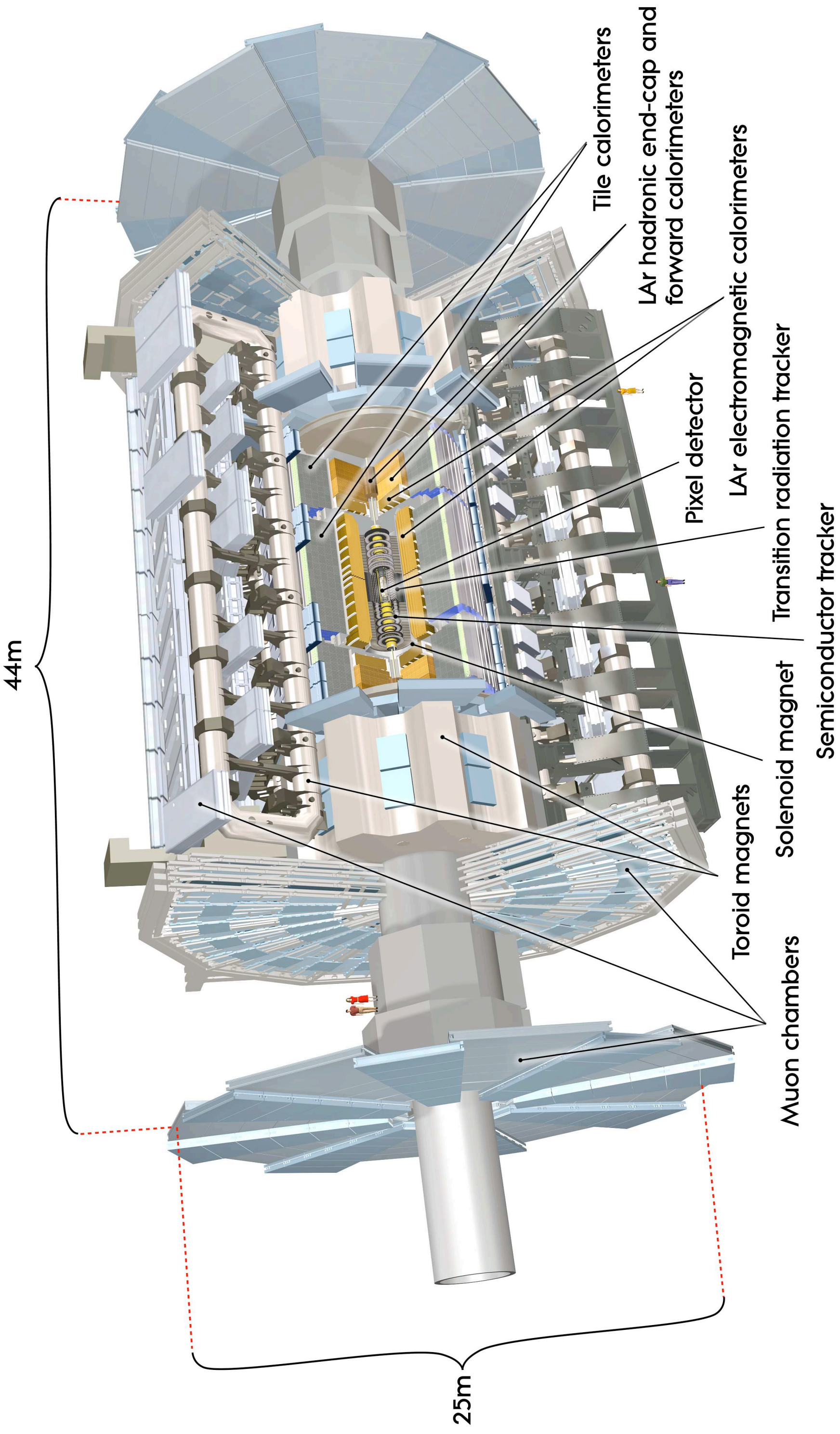


Figure 2.2: The ATLAS detector

Level 1, Level 2 and the Event Filter (L1, L2 and EF), which in three steps decide whether an event should be kept for further analysis. The staged approach of this trigger system allows each consecutive step to take more time, and process more information before making a decision. Only a small fraction of events will contain interesting information. These events are further processed by the DAQ system. All other events are discarded.

In the following sections the layout, construction and goal of each detector will be described in more detail. More information can be found in the ATLAS detector paper[24].

2.3 Inner Detector

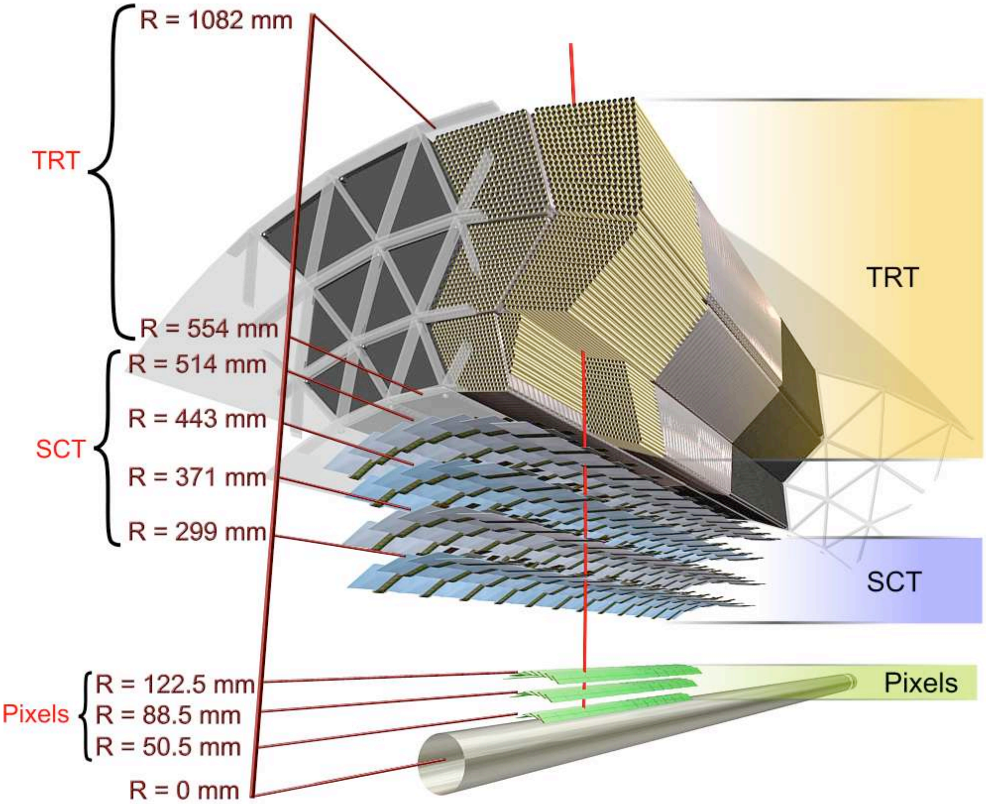


Figure 2.3: Cutaway view of the Inner detector barrel.

The Inner Detector (ID) of ATLAS, shown in Figures 2.3 and 2.4, is designed to provide tracking information for charged particles, and to determine the location of primary and secondary vertices in the event. The ID consists of three subsystems: the pixel detector for vertex location, the Semi Conductor Tracker (SCT) for high resolution space points near the interaction, and the Transition Radiation Tracker (TRT) for a long measuring arm for higher p_T resolution. The Inner Detector covers $|\eta| < 2.5$ for p_T measurement, and provides electron

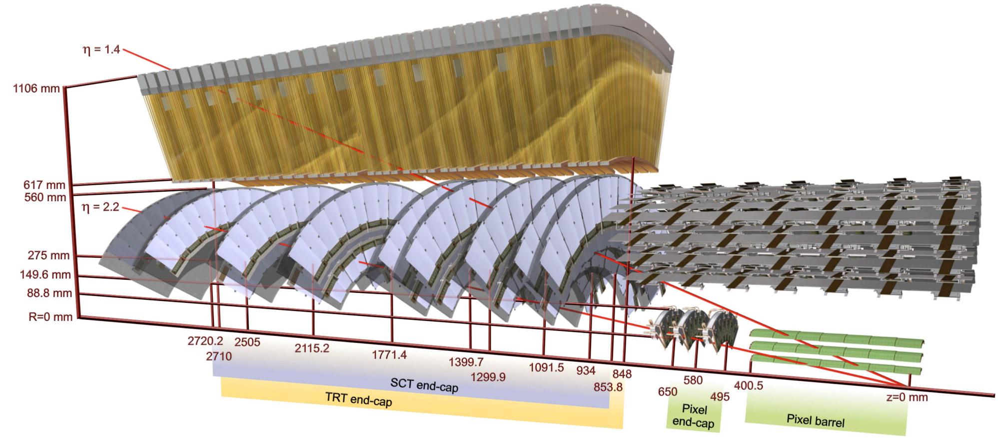


Figure 2.4: Cutaway view of the Inner detector endcap.

identification with the TRT up to $|\eta| < 2$. The three subsystems of the ID will be discussed in the following subsections.

2.3.1 Pixel detector

The pixel detector consists of three layers in the barrel (numbered 0-2), and three disks in each endcap. Figures 2.3 and 2.4 show the positions of the barrel and endcap modules. To provide the granularity needed in the busy events expected at ATLAS, the pixels have a size of $50 \times 400 - 600\,\mu\mathrm{m}^2$. The modules on which the pixel sensors are mounted are the same in the barrel and the endcap region. Each module is a layered structure formed by the front-end electronics chips (16 per module), bump-bonded to the pixel sensors, glued to a flexible printed circuit board. This board also contains the Module Control Chip. The total active area of the pixel detector is approximately 1.7 m^2.

The high levels of radiation which the ATLAS detector systems will be subjected to place more constraints on the detector design. Since it is closest to the beam pipe, the pixel detector's layer-0 (the so-called b-layer) will suffer the highest radiation fluxes. As a consequence, it is expected to need replacement after three years of LHC operation at design luminosity.

One effect of radiation is that the n-type silicon will become p-type after a neutron equivalent fluence $\mathrm{F}_{\mathrm{neq}}$ of $\sim 2 \times 10^{13}\,\mathrm{cm}^{-2}$. After type inversion, the effective doping will increase over time, which in turn will cause an increase in the depletion voltage and leakage current . To keep voltage and current values within acceptable limits, both the pixel detector and the SCT need to be operated at a temperature of -5 to -10 °C. To maintain good charge-collecting efficiency after type inversion, the pixel sensors are made using an n/n$^+$ junction, instead of the more traditional, and more easily made, p/n junction. As a last measure, highly oxygenated silicon, which has a demonstrated high radiation tolerance, has been used to construct the sensors.

The alignment of the pixel detector is critical to achieve the vertex location precision needed at ATLAS. To achieve this alignment the pixel sensors are mounted on staves which are placed in a carbon fiber frame. Five mount points assure that the staves are kept in position. This frame is positioned inside the SCT with 100 µm accuracy.

2.3.2 Semiconductor Tracker

Like the pixel detector, the Semi Conductor Tracker (SCT) provides high resolution tracking and pattern recognition. See Figures 2.3 and 2.4 for the position of the SCT barrel and endcap modules. At the larger radii, pixels are no longer required to obtain the desired granularity. Thus the SCT strip detector sensors contain single sided 12 cm long p-in-n strips, with a average pitch of 80 µm. The SCT modules have two sensors glued back to back with a 20 mrad stereo angle between the sensors. The modules are designed to work at an operating bias voltage over the p-n junction of 150 V, gradually increasing with increasing radiation dose to 350 V. The final depletion voltage will depend on both total integrated luminosity.

2.3.3 Transition Radiation Tracker

The Transition Radiation Tracker (TRT) serves two functions. Because of its large radius it can give the most accurate measurement of the momentum of charged particles passing through the ID. For this reason the location of a traversing particle is only measured in the bending direction. This measurement has to be combined with the 3D measurement of the pixel and SCT detectors. The second function is separation of electrons from charged pions, using transition radiation emitted by electrons when traversing the TRT.

The TRT consists of drift tubes (straws) of 4 mm in diameter. The stability of the straws is provided by carbon fibers. The length of the straws is 144 cm in the barrel and 37 cm in the endcap. In between the straws are the transition radiation fibers that produce photons when an electron passes through. The anodes of the tube are gold-plated tungsten wires. The tubes are filled with a gas mixture of 70% Xe, 27% CO_2 and 3% O_2. A high voltage between the tube and the wire induce an avalanche when an ionizing particle traverse the tube, thus creating a signal.

The transition radiation photons produced by electrons traversing the TRT are absorbed by the Xe gas, and produce a much higher signal amplitude than that of minimum ionizing particles traversing the detector. The difference between signals from tracking and from transition radiation is determined separately for each straw, using a low and a high signal threshold.

In the barrel, the wires are separated in the middle, to reduce the occupancy in the TRT. In the inner nine layers the occupancy is so high that the wires are cut into three parts, where only the outer parts are connected to the readout.

A TRT barrel module consists of a carbon fiber shell, containing an array of straws and transition radiation fibers. Inside the module shell CO_2 gas flows to prevent discharges and to flush out any leaking Xe, which would otherwise invisibly absorb transition radiation photons. A HV plate is located at each end of the module, which connects the walls of the tubes to the HV. Mounted on these plates are the front end electronics boards connected to the wires of the tubes.

The modules are combined in three layers to form a TRT barrel wheel. Figure 2.5 shows one quarter of a TRT wheel, with an inner, middle and outer module highlighted. Each wheel contains 73 layers of straws interleaved with fibers. Three wheels together form the TRT barrel.

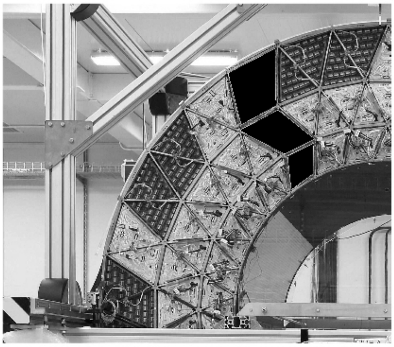


Figure 2.5: Quarter of a TRT wheel with three modules (inner, middle and outer) indicated in black.

The TRT endcap consists of two separate parts. The one closest to the interaction point contains 12 wheels with eight layers of tubes. The outer part contains eight such wheels. Each layer in a wheel contains 768 radially oriented straws. Each consecutive layer is rotated with 3/8 of the azimuthal spacing within a layer. In total, there are 160 straw planes interleaved with transition radiation foil in both endcaps. At least 36 straws will be traversed by all tracks with $|\eta| < 2$, with the exception of tracks with $0.8 < |\eta| < 1$, where the minimum number of traversed straws is 22.

2.4 Calorimeter

The calorimeter system, shown in Figure 2.6, is used to provide precise energy measurement for photons, electrons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the region $|\eta| < 4.9$. Furthermore the calorimeter information is used to distinguish electrons from photons up to $|\eta| < 2.5$ (limited by the ID). The different subsystems of the calorimeter will be described in the following sections.

2.4.1 Liquid Argon Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter consists of a barrel part covering $|\eta| < 1.475$ and an endcap for $1.375 < |\eta| < 3.2$. The calorimeter is a lead-Liquid Argon (LAr) detector. The lead absorber plates and kapton electrodes are folded in an accordion shape, providing full, homogeneous azimuthal covering. The LAr serves as active material. The bulk of the EM calorimeter is located on the outside of the solenoid magnet which provides bending power for the inner tracker. To take into account for the extra material introduced by the solenoid, a LAr presampler is placed in front of the magnet. Magnet and calorimeter are placed in the same cryostat.

The barrel calorimeter is split at z=0 into two identical halves, separated by a 4 mm gap. The endcaps both have their own cryostats, which they share with the Hadronic End-cap Calorimeter (HEC) and the Forward Calorimeter (FCAL).

The readout of the EM calorimeter is split in separate readout towers in η and ϕ. Each tower is split in three layers in R, with the inner layer having the finest granularity, to precisely pinpoint the start of the shower. This is shown in Figure 2.7.

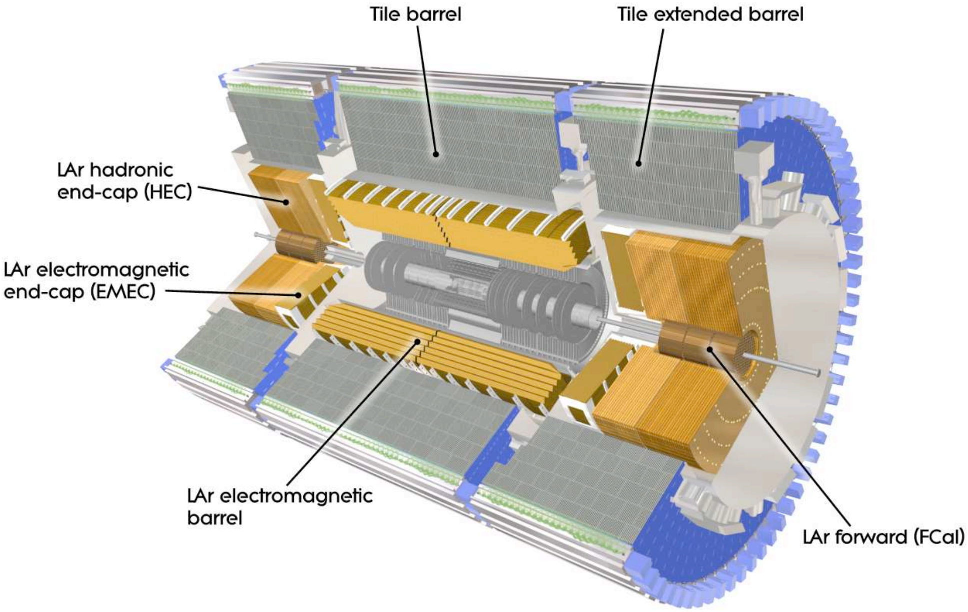


Figure 2.6: Cut-away view of the ATLAS calorimeter.

The total thickness of the EM calorimeter is at least 22 X_0 in the barrel and at least 24 X_0 in the endcap, where the radiation length X_0 is the mean distance over which an electron looses $1/e$ of its energy through electromagnetic interactions.

2.4.2 Tile Calorimeter

The Tile Calorimeter is the ATLAS hadronic calorimeter in the barrel. It has steel absorber plates interleaved with scintillator plates as the active material. It is segmented radially in three layers, with coarser granularity in η with increasing radius, and in the azimuthal direction in 64 modules. In η, the readout cells are made pseudo-projective towards the interaction region. The total thickness of the detector at the outside of the Tile Calorimeter is 9.7 λ, where λ is the nuclear interaction length, the mean distance over which a hadronic particle looses $1/e$ of its energy through nuclear interactions. The Tile Calorimeter covers the range $|\eta| < 1$ for the barrel and $0.8 < |\eta| < 1.7$ for the extended barrel, see Figure 2.6.

2.4.3 Liquid Argon Hadronic Endcap Calorimeter

The HEC is a LAr/copper plate sampling calorimeter, covering $1.5 < |\eta| < 3.2$, overlapping both the extended barrel at lower and the FCAL at higher η. Thus the HEC provides a smooth transition with respect to the amount of material 'seen' by particles passing through the transition regions between the different calorimeter subsystems. The HEC consists of two wheels, each separated longitudinally in two layers. In phi, the wheel is divided in 32 segments.

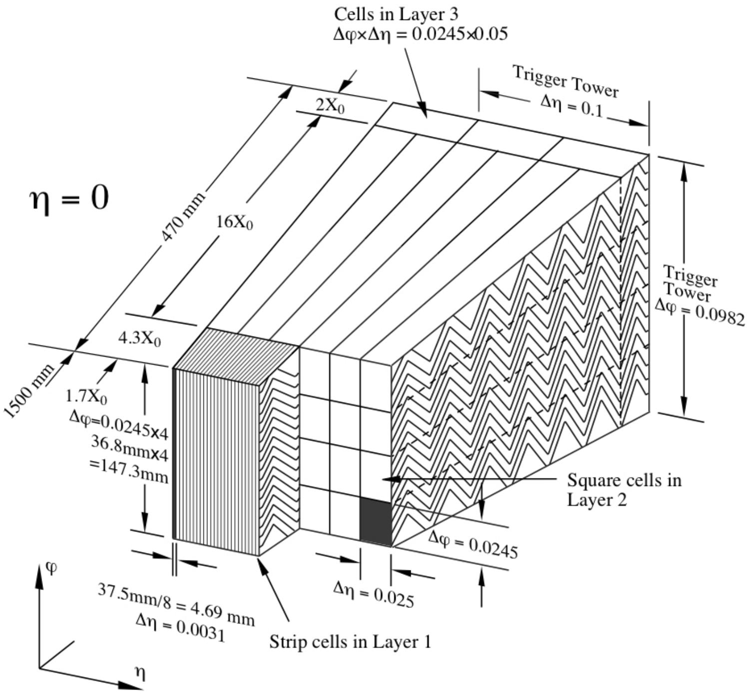


Figure 2.7: The granularity of the LAr barrel calorimeter.

The cells of the HEC have a size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for the inner layers, and 0.2×0.2 for the outer layers.

The absorber material is made in plates of thickness from 25 to 50 mm, getting thicker further away from the interaction point. The LAr gaps are 8.5 mm wide.

2.4.4 Forward Calorimeter

The FCAL is housed in the same cryostat as the endcap LAr calorimeters. It is the system that closes the $|\eta| < 4.9$ coverage that was required to provide good $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement for the ATLAS detector. To reduce the neutron background from interactions with the calorimeter in this high-radiation environment, the FCAL front is placed about 1.2 m further back with respect to the other endcap calorimeters. To keep the number of interaction lengths at 10, the density of the FCAL must be increased.

The FCAL is divided in three parts. The first part, with copper as the absorber, is made for electromagnetic measurement, and the other two, which have tungsten as the absorber, form the hadronic part of the FCAL. LAr is the active material in all three parts. The absorber material is made in rods placed inside tubes, hold in place by a solid metal frame. This geometry allows for very small gaps (down to 0.25 mm), which are needed to avoid ion buildup in the active region in this high-radiation environment.

2.5 Muon System

The outer layers of ATLAS are formed by the Muon Spectrometer with the toroid magnets providing bending power. The spectrometer uses four different technologies for different purposes (triggering, precision measurement), and different locations (central, forward in η). The location of the four different types of chambers are indicated in Figure 2.8. The Monitored Drift Tubes (MDT) provide precision measurement in the bending (η) direction in the barrel and a large part of the endcap, over the full η coverage of the spectrometer ($0 < |\eta| < 2.7$). In the inner layers of the endcap, where radiation levels and occupancy would be too large for the MDT, the precision tracking is provided by the Cathode Strip Chambers (CSC).

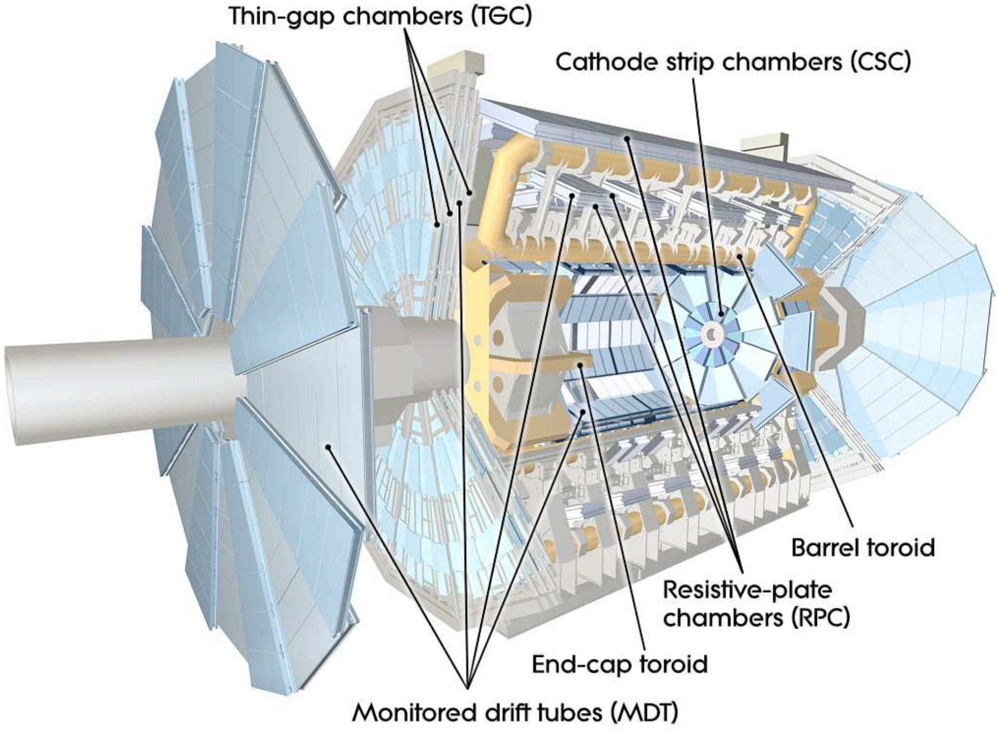


Figure 2.8: The muon spectrometer system.

The other two types of chambers, the Thin Gap Chambers (TGC) and the Resistive Plate Chambers (RPC), fulfill two functions: they provide a fast signal, which is used for triggering and for bunch crossing identification, and they provide a measurement of the second coordinate perpendicular to the bending direction.

The magnetic field in the barrel ($|\eta| < 1.4$) is provided by the large air-core magnet toroid. In the endcaps ($1.6 < |\eta| < 2.7$), an endcap toroid provides the magnetic field, and in the transition region ($1.4 < |\eta| < 1.6$) a combination of barrel and endcap toroid magnetic field ensures enough bending power for the muons. The bending power provided by the magnets is 1.5 to 5.5 Tm in the barrel, and 1 to 7.5 Tm in the endcap region, while being somewhat lower in the transition region.

2.5.1 Monitored Drift Tubes

The tubes of the MDT are made from 0.4 mm thick aluminum, with an outer diameter of 29.970 mm. Muons passing through will ionize the Ar/CO_2 gas in the tube. The gas is kept at a 3 bar absolute pressure. The electrons from this ionization will drift onto the 50 µm thick tungsten/rhenium wire. A high voltage between the wire and the tube wall induces an avalanche close to the wire which amplifies the signal. The tube measures the drift time of the electrons, thus providing a measure of the distance of closest approach of the muon track to the wire. The maximum drift time is 700 ns, thus many bunch crossings. This makes the bunch crossing identification by the trigger chambers important. The trigger chambers measure the location in η and ϕ, and provide a timestamp for bunch crossing identification. The trigger chamber hits are matched to hits in the MDT, which then provides a more precise measurement in the η direction. If there is more than one hit in a single tube in one event, this matching will fail. The probability of having more than one track through a single tube is estimated to be about 2×10^{-6}, if the tracks are uncorrelated. Correlated close-by tracks need to be matched to the ID to be resolved.

As seen in Figures 2.10 and 2.11 the MDTs are placed in three concentric shells in the barrels, and three disks in the endcaps. In the barrel, the chambers are mounted within and on the barrel toroid magnet. In the endcap, there is one layer in front of, and two layers after the endcap toroid. The chambers layout follows the eight-fold symmetry of the toroid magnet, with each segment of the detector filled by two sectors, one with a large, and one with a small chamber, see Figure 2.11. As can be seen in the same figure, the small and large chambers overlap at the edges, to prevent gaps in the sensitive area of the spectrometer. Some gaps cannot be avoided. There are four gaps in sectors 12 and 14 where the feet of the ATLAS support structure are placed. There are also some gaps at $\eta = 0$ to allow services for the calorimeters and ID to pass. This means that high momentum tracks are not recorded in $\eta < 0.08$ for the large and $\eta < 0.04$ for the small chambers.

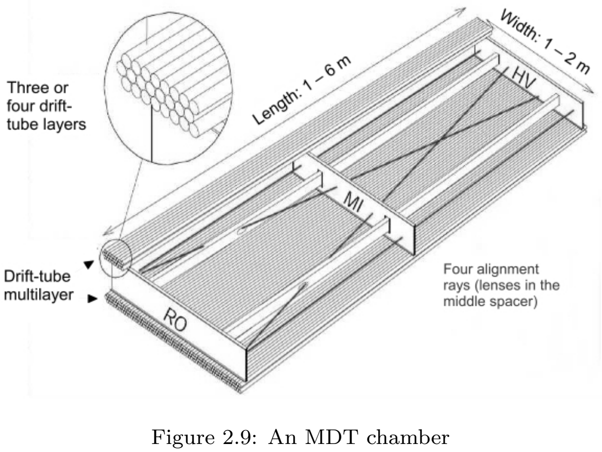


Figure 2.9: An MDT chamber

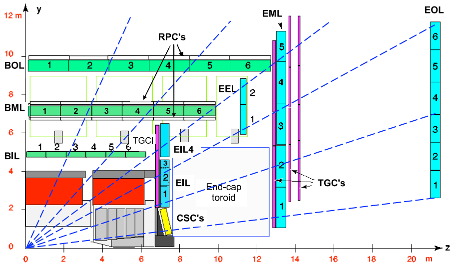


Figure 2.10: Cross section along the beam axis of the ATLAS muon system

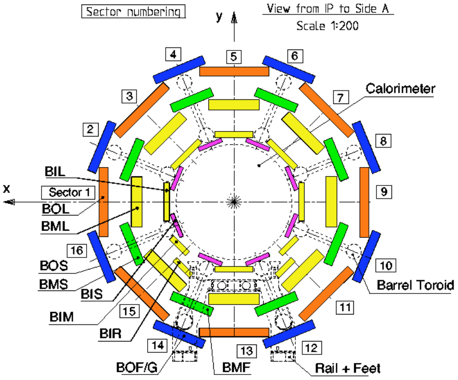


Figure 2.11: Cross section in the (x, y) plane of the ATLAS muon system

One MDT chamber has two multilayers, containing three layers of tubes for the outer and middle detector shells, and four tube layers for the inner shell of the detector. The two multilayers are separated by a spacer, which has a width 6.5, 170 and 317 mm for the inner, middle and outer shells of the detector. Thus each chamber provides a track segment that can be matched with segments from chambers in the other layers. The spacer structure also provides the mounting points for the HV supplies, the readout electronics and B-field sensors. Finally, each chamber has an optical alignment system that monitors deformations of the chamber due to thermal variations and magnetic field. Figure 2.9 shows a schematic view of a MDT barrel chamber.

To achieve the required muon p_T precision (10% for a 1 TeV track standalone (without ID)), very stringent constraints have been placed on the construction and positioning precision. The spectrometer standalone resolution is given by $\delta p/p = \Delta S/500\ \mu m \times p/TeV$, with ΔS the resolution of the track sagitta, which is given by $\Delta S = 45\,\mu\text{m}$ for three shells crossed by a track.

2.5.2 Cathode Strip Chambers

In the very forward regions ($2 < |\eta| < 2.7$) in the inner layer of the spectrometer, the hit rate in the MDTs would exceed the level where safe and durable operation can be guaranteed. In this region the MDTs are replaced by Cathode Strip Chambers (CSC). These are multi-wire proportional chambers, which are read out through segmented cathodes. The wires of these chambers are oriented in the radial direction. Perpendicular to the wires are the finely segmented cathode strips that provide the precision measurement in the bending plane. In the second cathode plane, the strips are oriented parallel to the wires, providing the second coordinate measurement, and a fast signal for bunch crossing identification. Reading out the cathode strips instead of the anode wires increases the spatial resolution of the CSCs. One CSC chamber has four active layers.

The CSCs are placed in two separate disks, see Figure 2.12, one disk with eight large CSC chambers, and one with eight small chambers. The disks are placed such that the chambers of one disk fill the gap between the chambers of the other, with a small overlap between the

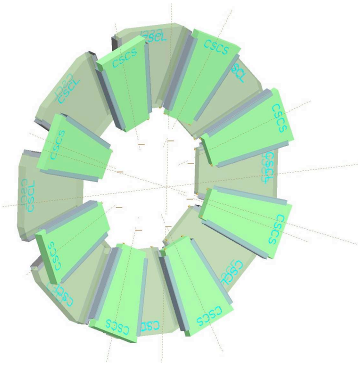


Figure 2.12: The CSC disks

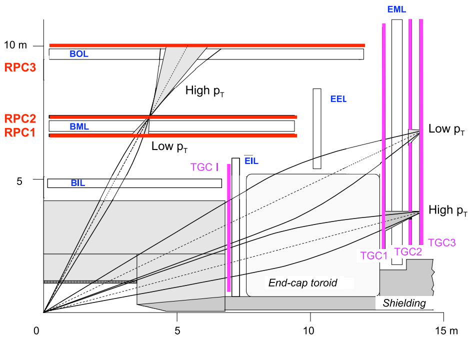


Figure 2.13: The location of the RPC and TGC chambers, together with the location of the pivot planes

chambers. The chambers are tilted by 5.3° towards the interaction point. This is to make sure that the tracks through the chambers are perpendicular to the chamber plates. Inclined tracks would reduce the spatial resolution of the chambers.

The CSC achieve a resolution per layer of 60 µm in the bending plane, and about 5 mm in the non-bending plane. The alignment of the CSCs uses the same system of optical alignment and alignment bars as the MDT.

2.5.3 Resistive Plate Chambers

The MDTs and CSCs provide a precise measurement of the coordinate of a track in the bending direction, which results in a precise measurement of muon p_T. What is missing is the measurement linking a muon hit to a specific bunch crossing, the measurement of the second coordinate (although the CSC can do both), and the ability to trigger on a muon with a certain p_T. These measurements are performed by the Thin Gap Chambers (TGC, see next section) in the endcap, and the Resistive Plate Chambers (RPC) in the barrel.

The RPCs are mounted on the MDT chambers in three concentric circles, the two inner layers around the MDT middle layer, the outer layer on the inside or outside (as seen from the interaction point) of the MDT outer small or large chambers respectively, see Figure 2.13. The L1 trigger requires a coincidence in the inner and middle layer for the low-p_T trigger and a third coincidence in the outer layer for the high p_T trigger. To apply the p_T threshold, the slope of the track from the pivot plane (the middle layer) to the interaction point is compared to the slope of the track to the coincidence planes. This is implemented by allowing the coincidence within a certain envelope centered on the pivot hit for the coincidence hit(s) (see Figure 2.13). Requiring the high p_T trigger events to first pass the low-p_T threshold helps reduce the trigger fake rate.

Each RPC chamber has two active gas volumes, both of which are read out by two plates with orthogonal strips measuring both coordinates. The gas volume is delimited by two resistive plates, which are kept at HV and ground potential by two graphite electrodes. The ionized track of a particle passing through the detector causes an avalanche in the high potential gradient (4.9 kV/m). This avalanche induces a charge on the pick-up strips, which are read out. The signal is sent to coincidence units near the detector, where the p_T thresholds are applied. The result of the coincidence is sent to the Central Trigger Processor (CTP).

2.5.4 Thin Gap Chambers

A muon in the endcap with a given p_T has a much larger momentum than a muon with the same p_T in the barrel. This means that for a muon to pass a certain p_T threshold, the bending of the track is much less in the endcap than in the barrel, which, together with the higher rate, calls for a different technology and a finer granularity. To this end, the Thin Gap Chamber was chosen. This is a multi-wire proportional chamber like the CSC. The gap between the wires and the cathodes (1.4 mm) is smaller than the gap between the wires (1.8 mm). The small gaps give a short drift time and thus fast response needed for the trigger system. The chambers are placed such that the angle between the muon track from the IP and the chamber normal is always larger than 10°. This means that there are no tracks that pass exactly perpendicular between two wires, where the field gradient is small, causing long drift times. The wires are read out to provide the coordinate measurement in the η direction, while cathode strips provide the measurement of the second coordinate.

The TGCs are mounted in two layers near the CSC/MDT inner layer of the endcap. Seven TGC layers are mounted on the middle layer of the MDT using two doublet and one triplet unit, consisting of two and three gas volumes with electrodes respectively, separated by a paper honeycomb.

The Front End (FE) electronics on the chambers asks for a coincidence in three out of four layers of the two doublets or two out of three layers of the triplet modules. If this is found, the data is sent to a coincidence matrix which looks for patterns corresponding to muon tracks above threshold.

2.5.5 The MDT read out system

Each subsystem of ATLAS has a dedicated readout system in place, designed to obtain the information from the detector, deliver it to the Central Trigger Processor (CTP) if applicable, and when a L1 accept arrives, package it in a predefined ATLAS data format, such that event fragments can be stored on the Read Out Buffers (ROB). These fragments are read out by the L2 trigger (see section 2.6), and when accepted, combined by the event builder. One example of a subsystem-specific readout chain will be discussed here, namely that of the MDT system.

2.5.6 Front-end electronics

The front-end electronics system consists of three parts. Up to 24 channels are connected to a single mezzanine card. This card houses three Amplifier/Shaper/Discriminator (ASD) chips, which measure pulse hight. The signal from these chips are send to a single Time to Digital Converter (TDC) on the same card which measures the arrival time of the incoming signals, and stores the result in a buffer. When a L1 accept signal arrives, up to 18 TDCs on one chamber deliver this data to a Chamber Service Module (CSM). The TDC and CSM are mounted on the chambers. Each TDC adds a header and trailer to its data stream to the CSM, containing the TDC id, event id, and bunch crossing id. The CSM deserializes and multiplexes the data from the TDCs, and sends them through an optical link to the Muon Readout Driver (MROD).

2.5.7 MROD

The MROD module receives the input from up to six CSMs, and packages them to form an event fragment. It also has processors that run error checking algorithms. The format of the data delivered by the MRODs to the ATLAS readout systems is shown in Figure 2.14. Figure 2.15 shows a schematic of the MDT readout system. The MRODs are located in a separate cavern, behind a shielding wall. Thus the MROD design does not have to take radiation hardness in to account, which enables the design to be more complex than for the on-detector electronics.

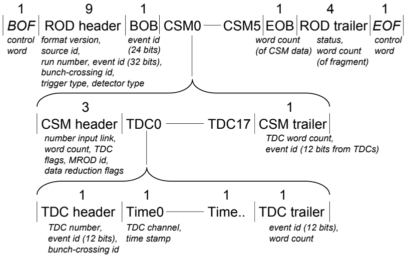


Figure 2.14: The MDT data format.

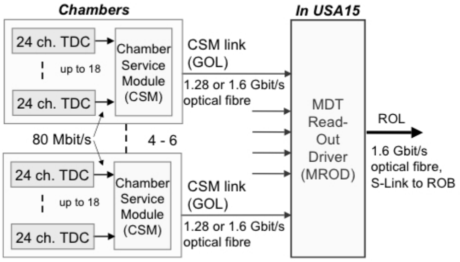


Figure 2.15: Schematic of the MDT readout system.

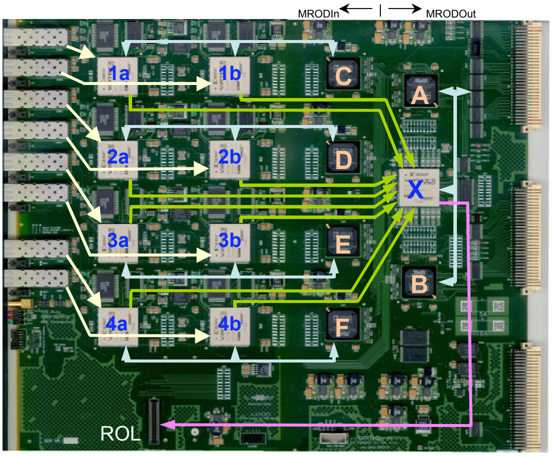


Figure 2.16: The MROD X module. Shown here are the two FPGAs (a and b) of the MRODin 1-4, the MRODin SHARC DSPs C-F, the MRODout SHARC A and B, and the MRODout FPGA X. Also shown are the RocketIO links between MRODin and MRODout FPGA, and the communication links between the SHARCs and the FPGAs. Not shown are the SHARC links between the DSPs. Each MRODin DSP is connected with each MRODout DSP, and with its neighboring MRODin DSP. DSP A and B are also connected to each other.

Figure 2.16 shows a picture of the final MROD module, the MROD-X, which is installed in the ATLAS read-out system. The main functional components of the MROD are indicated in this picture. On the left-top side are four identical MRODin modules, containing an optical input, two Field Programmable Gate Arrays (FPGAs) which collect and format the data, and a SHARC Digital Signal Processor (DSP) which monitors the data flow, and resolves detected errors. On the right side of the module lies the MRODout, with one FPGA for event fragment building, and two DSPs for monitoring and steering. From the MRODout the data is delivered to the optical ReadOut Link (ROL).

2.6 Trigger, Data Acquisition and Detector Control

With the 25 ns time between bunch crossings delivered by the LHC, ATLAS will take data at a rate of 40 MHz. The trigger system is designed to filter out most of the uninteresting (minimum bias) events, while being very efficient in keeping those events that ATLAS was designed to investigate. Uninteresting here is a subjective term, that may change over the years of running ATLAS, when understanding of the physics at the LHC increases. Thus a flexible system is needed, which can accommodate the unexpected. The trigger system is divided in three levels. The first level (L1) is a hardware based trigger system. It is designed to bring the event rate down to 75 kHz. Level 2 and the Event Filter consist of computer farms, and bring the total event rate down to about 2.5 kHz and then to the final rate 200Hz, with an average event size of 1.3 MB.

The data acquisition (DAQ) system consists of all the elements that transport the data of the detector, keep data from those events that pass the trigger requirements, collect all subdetector data from the same bunch crossing and store it on tape.

The Detector Control System (DCS) allows operators to monitor the state of the detector and turn on/off the different subsystems or change operating parameters. The DCS allows safe operation of the detector. Furthermore, it facilitates the communication with the LHC machine.

2.6.1 L1 Trigger

The L1 trigger is based only on information from the calorimeter and muon spectrometer, as can be seen in Figure 2.17. The Central Trigger Processor (CTP) takes information from all calorimeter systems, (EM, Tile, HEC, FCAL), but at a coarser granularity compared to the offline reconstruction. With this information it looks for indications of jets, electrons, photons or hadronic taus with large transverse energy E_{T} ($= E\sin\theta$), for events with a large total E_{T} summed over all calorimeter cells and for events with a large $E_{\mathrm{T}}^{\mathrm{miss}}$, the vector sum of all transverse energy measured in the detector. For the jets, there is also the possibility to cut on a high $\sum E_{\mathrm{T}}$ over all jets. An isolation requirement (minimum distance to other high-E_{T} objects) can be placed on the electrons, photons and taus. The CTP cuts on the number of (programmable) threshold crossings per object per bunch crossing.

The L1 muon trigger gets its information from dedicated trigger chambers (The TGCs in the endcap and the RPCs in barrel) in the muon spectrometer. The trigger is based on coincidence in stations within a road. The road is defined as the path of a muon through the spectrometer. The p_T threshold is implemented as the allowed width of the path: high p_T muon tracks are bent less by the magnetic field, and the allowed path is narrow. Low p_T tracks are bent more, and a wider path is allowed for the low p_T threshold. The information from the trigger chamber is sent to dedicated hardware that serves as an interface between

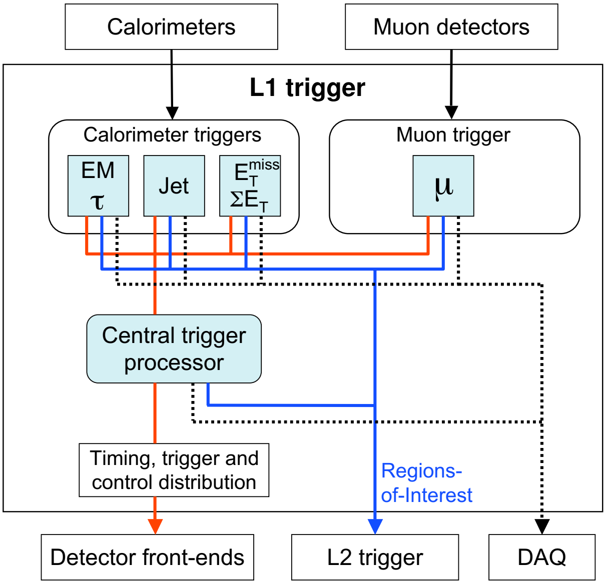


Figure 2.17: Block diagram of the L1 Trigger system

the muon system and the CTP. This interface hardware combines the number of threshold crossings in barrel and endcap, taking care to remove double counting by overlapping trigger chamber. The interface hardware sends the number of threshold crossings on to the CTP. Six distinct thresholds for the muon p_T can be programmed, three low p_T (6-9 GeV) and three high p_T (9-25 GeV). The coincidences between the hits in the different layers within a path are determined by programmable custom electronics on (in the barrel and endcap) and near (endcap) the detector systems.

The L1 trigger decision is based solely on the number of times a certain threshold has been crossed. But the L1 trigger system also passes on the location in η and ϕ where this occurred. This information is used to built Regions of Interest (RoIs), which are used as seeds for the L2 trigger processing, where the full detector information is available, within the RoIs. This drastically reduces the amount of data that is needed for the L2 processing.

The CTP uses the information of the number of threshold crossings in each trigger condition, plus the flags for events exceeding a ΣE_T or $E_\mathrm{T}^\mathrm{miss}$ threshold. The list of trigger items can be up to 256 items long. These items also contain extra triggers, such as beam pick-up and random triggers. The CTP also receives the LHC clock signal, and through the Trigger Timing and Control chip distributes this clock to the all the detector systems.

2.6.2 High Level Trigger and Data Acquisition

The Data Acquisition system is responsible for getting the data selected by the L1 trigger from the detector through the High Level Trigger (HLT) to storage. A schematic overview of the DAQ/HLT is shown in Figure 2.18. When a signal is found in a sub detector, it is stored in buffers which are located on the front end electronics boards on or near the detector. Data from specific detectors is sent to the CTP as described above. When a L1 accept signal is received, the data is transferred from the detector front end to the Read Out Drivers (ROD), which package and format the data in a detector specific manner such that it can be stored in the Read Out Buffers (ROB), which are located in the ROS (Read Out System) PCs. In the meantime, the L1 RoI information is sent to the RoI builder. The output RoI structure is sent to the L2 supervisor and used by the L2 trigger as seed for the selection algorithms. The L2 trigger requests data through the supervisor from specific ROSs, typically containing data within an RoI, to apply the selection. This is done in a smart way, such that when it becomes clear that an RoI will not contribute to passing the L2, processing of that RoI is stopped.

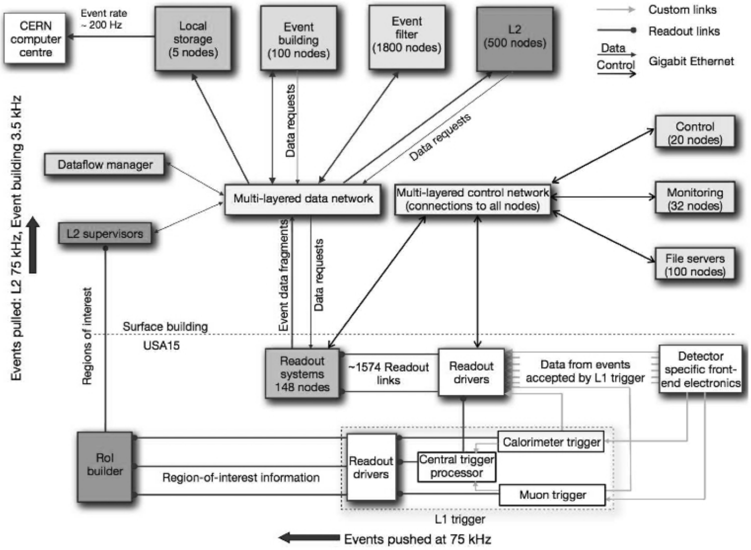


Figure 2.18: Schematic overview of the DAQ/HLT system

From here on, control is handed over to the dataflow manager (DFM), which takes the information from the L2 decision. When the L2 discards the event, the DFM sends a signal to all ROSs to expunge the data for that event. When an event is accepted, it requests all data fragments for this event from the ROSs, and sends it to the event builder, where it is combined in a single data structure. When the event is built and sent to the event filter (EF) for the final trigger selection, the DFM sends a message to all ROSs to expunge the data.

The EF consists of a computer farm, which uses offline reconstruction algorithms using information from the entire event. To reduce the time spent on each event, the EF is seeded on the results of the L2 trigger. When the EF accepts the event, it is sent to an output node of the DAQ, where the event is stored on a local disk. On this local storage, the events are divided in streams, based on the trigger selection criteria that were passed. One event may be placed in multiple streams. From these output nodes the events can be picked up and transferred to mass storage.

2.7 Reconstruction, storage and analysis on the Grid

With an event output rate of 200 Hz and an average event size of 1.3 MB the total size of the data stream coming from ATLAS alone is 260 MB/s, not counting events that are stored in multiple streams and detector status information. For all LHC experiments combined this results in several PetaBytes per year of data that has to be reconstructed, stored and analyzed. No single computing facility has the storage, computing and networking capacity to handle this much data. This is why the LHC Computing Grid (LCG, see their website [30]) was designed, a global network of computing facilities that together make sure that all data from ATLAS and the other LHC detectors is stored (more than once) and reconstructed. Furthermore, simulations of detector data are made and stored on the Grid. Finally users (physicists) run their analysis on the Grid, sending their job to data, instead of downloading the data to where the job is run. This causes a considerable reduction of the bandwidth load on the Grid.

Since there are so many users, admins, sites, pieces of hardware and software, types of jobs, the Grid has to be a very flexible environment that does not depend on a single site, user or machine to work as expected. In fact this is one of the strong points of the Grid: due to its large size any malfunction can be circumvented.

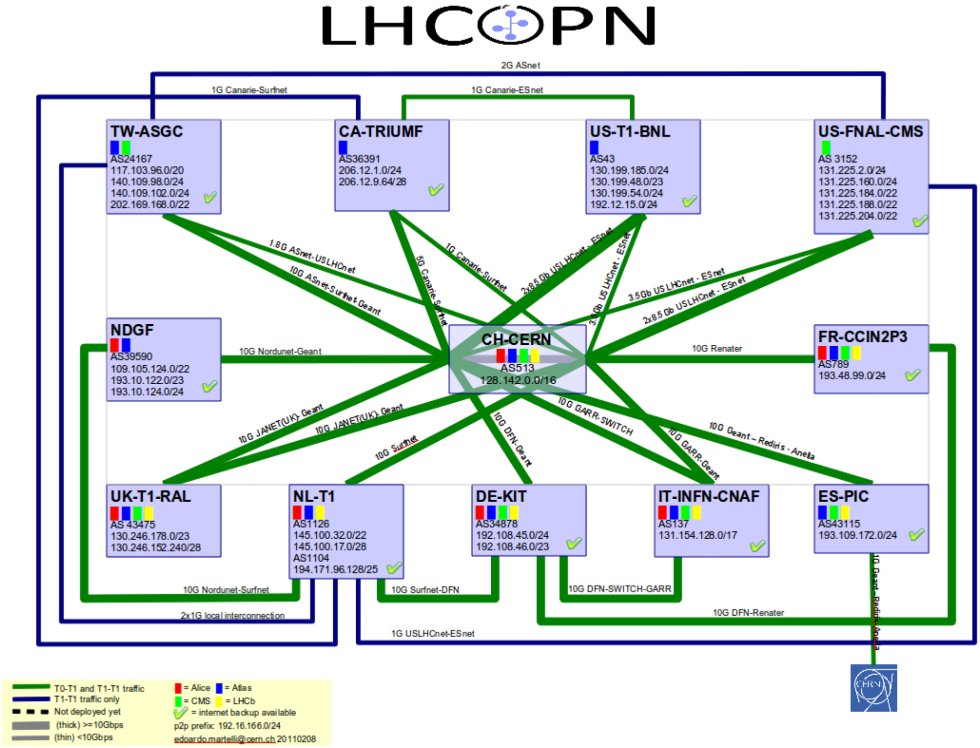


Figure 2.19: The Tier-0 and Tier-1s of the LHC Computing Grid.

To achieve this the Grid is divided in several layers both by location and by component. The different locations around the world are structured hierarchically in tiers. The first tier is a single Tier-0 at CERN, where all data originate. The Tier-0 stores all Raw data from the detectors, does a first reconstruction of the data and handles reprocessing of the data when the detectors are shut down. It passes the reconstructed and raw data on to the Tier-1s, of which there are eleven, distributed over Europe, Asia and North-America. See Figure 2.19 for a schematic overview of the connections between the Tier-0 and the Tier-1s. Each Tier-1 stores backup copies of (a part of) the raw and reconstructed data, and handles large scale reprocessing. They distribute this data to the next layer, the Tier-2s, and store the simulated data produced at the Tier-2s. Besides producing simulated data, the Tier-2s handle analysis jobs from users on data that is stored at or close by that Tier-2.

The Grid is also divided in four component layers. First there is the physics software, programs used to do data analysis. The second software layer, the middleware, deals with communication between physics software and the Grid sites, and takes care of user authentication. The third layer is the fabric layer, containing hardware (servers, storage) and the software needed to operate these. The fabric layer is managed automatically, to smoothly handle updates of both soft- and hardware. The last layer is that of data transfer, consisting of a dedicated optical network connecting CERN to each of the Tier-1s.

With this enormous computing power and storage capacity at his/her disposable, it is extremely easy for a user to get lost in the digital jungle that is the Grid. Good bookkeeping is essential, and there are tools available that can help. Ganga is one such tool. It has the advantage that a job is built modularly. This allows the user to run the same job at different location (on a PC, a local batch system, or on the Grid), which makes debugging of jobs much easier. Specific plug-ins have been made for Ganga that support for instance running jobs within the ATLAS reconstruction and analysis framework Athena. Ganga also makes it very easy to split a job into several subjobs, by running over different subsets of a dataset. Each subjob can be managed (killed, canceled, restarted) separately. Since the user has no control over the hard- and software on which a job is run, it is crucial that restarting a job that was killed (for instance because of a server shutdown) takes little effort. Finally, since the Ganga interface takes python commands it is very easy to automate job submission, which is invaluable when running e.g. the same job over many different (simulated) datasets.

Chapter 3

Identification of physics objects in the ATLAS detector

This chapter will describe the particle identification criteria used in ATLAS. The definitions used for the analysis of chapter 6 were fixed in early 2008, before the first collision data had been collected. With the first measurements of collision data in september 2008, fall 2009 and early 2010, understanding of the detector increased significantly between the start of the analysis of chapter 6 and the analysis of chapter 7. Furthermore the development of the ATLAS analysis framework progressed, and with it the possibilities of defining and understanding the reconstructed physics objects. The next sections deal with the definition and measurement of leptons, jets and missing transverse energy as used for both the 2008 analysis on MC, and the 2010 analysis on data. Where the definition of an object differs for the two analyses presented in this thesis, both definitions will be discussed. Some definitions for the 2008 MC analysis are identical to those used in the ATLAS publication on the expected performance of the ATLAS detector [31].

3.1 Jets

Analysis of 2008 MC simulation

Several jet reconstruction approaches and definitions are available in ATLAS which meet theoretical requirements of infrared and collinear safety and order independence, as well as following the experimental guidelines of detector technology and environment independence. For a long time however these algorithms were computationally too time-consuming. Therefore the jet definition originally in use by ATLAS, and used in the analysis of chapter 6 uses a fixed-sized seeded cone algorithm, which is not infrared and collinear safe. It was was the standard in the SUSY working group of the collaboration at the time of production of the simulation samples. Specifically a narrow cone size of 0.4 was chosen because of the large multiplicity of jets in SUSY events.

ATLAS implemented an iterative seeded fixed-cone jet finder. First, all input is ordered in decreasing order in transverse momentum, p_T. If the object with the highest p_T is above the seed threshold and has at least $p_T > 1\,\mathrm{GeV}$, all objects within a cone in pseudorapidity

η and azimuth ϕ with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < R_{cone}$, where R_{cone} is the fixed cone radius, are combined with the seed. A new direction is calculated from the four-momenta inside the initial cone and a new cone is centered around it. Objects are then (re-)collected in this new cone, and again the direction is updated. This process is re-iterated until the direction of the cone does not change anymore, at which point the cone is considered stable and is called a *jet*. At this point the next seed is taken from the input list and a new cone jet is formed with the same iterative procedure. The jets found this way can share constituents, and a split-merge procedure is implemented in ATLAS. Jets which share constituents with more than half of the fraction of the p_T of the less energetic jet are merged, while they are split if the amount of shared p_T is below half of the fraction.

The input for the jet finder algorithm are topological cell clusters, for short *TopoClusters*. The clustering starts with seed cells with a signal-to-noise ratio above a certain threshold, $|E_{cell}/\sigma_{cell,noise}| > 4$. All directly neighboring cells of these seed cells, in all three dimensions, are collected into the cluster. Subsequently the cluster is extended with all neighboring cells that cross a lower threshold $|E_{cell}/\sigma_{cell,noise}| > 2$. Finally, a ring of guard cells without a specific threshold is added to the cluster.

The jets found, as described above, are constructed from the raw signal from the calorimeter cells. Since the ATLAS calorimeter is non-compensating, this raw signal has to be calibrated, to account for the difference in electromagnetic and hadronic response. In ATLAS a calibration scheme for calorimeter jets is applied based on cell signal weighting, called *H1* calibration after the experiment that helped develop and refine this approach [32]. To each cell a weight w is applied, which is a function of its location and its signal density, that is defined as electromagnetic energy signal divided by the cell volume. The weighting factor is ~ 1 for high density signals assumed to come from electromagnetic showers, and rising up to 1.5, the typical e/π signal ratio for the ATLAS calorimeters, with decreasing cell signal densities associated to hadronic showers. The weighting factor is applied to both the energy and the momentum terms of the cell four-momentum. All the calibrated cells of the jet are then summed up into a calibrated jet.

More refined corrections are needed to calibrate the jets to the particle level. Those include jet energy scale calibration, corrections for residual non-linearities in the jet response due to algorithm effects and suppression of signal contributions from the underlying event and/or pileup. Most of these can only be addressed in the context of a specific physics analysis after enough collision data has been acquired.

Analysis of 2010 data

With the coming of FastJet [33] algorithms it became possible to use the more elaborate algorithms which do meet the theoretical demands of infrared and collinear safety. For the analysis of chapter 7, a cluster algorithm is used known as the anti-k_T algorithm. The following two quantities are calculated for all calorimeter objects, which in this case are the topographic clusters:

$$d_{ij} = min(k_{Ti}^{-2}, k_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \tag{3.1}$$

$$d_{iB} = k_T^{-2} \tag{3.2}$$

$$\tag{3.3}$$

with ΔR in this case the distance between two objects in the (y, ϕ) plane, y being the rapidity of the object:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right). \tag{3.4}$$

What is called k_T in this algorithm is the transverse momentum of the object (p_T elsewhere in this thesis). The outcome of these two formulas are listed for all objects and combinations of two objects. If the smallest of this list is a d_{ij}, objects i and j are merged. If it is a d_{iB}, object i is defined as a jet, and removed from the list of objects.

This algorithm adds all soft objects within distance $\Delta R < R$ to a harder object, in order of closeness in ΔR. This ensures infrared and collinear safety of the algorithm. Nearby jets will automatically merge with clusters in between according to relative k_T and distance, so no split/merge algorithm is needed.

The jets are calibrated using EMJES calibration, which means that the correction for the hadronic response of the calorimeter is taken from Monte Carlo simulation.

Not all reconstructed jets are associated with a hard *pp* scatter. Some are due to atmospheric muons from cosmic ray interactions, from hardware issues such as calorimeter noise, or from beam-gas interactions or beam halo muons. Such ‘bad jets’ are rejected by application of a series of quality criteria.

If the relative fraction of total jet energy deposited in the EM calorimeter $EMF > 0.95$ and the jet quality $Q > 0.8$ the jet is assumed to be caused by noise in the EM calorimeter and tagged as bad. The jet quality is a measure of the difference between the measured and expected pulse shape in the calorimeter, $Q = \Sigma(a_i^{meas} - a_i^{exp})^2$. The quantities a_i are the sampled pulse heights as measured by the calorimeter, and the sum is over all samplings.

If the bad jet is caused by a noise in the Hadronic End-cap, this will show as a very narrow object forming the jet. Hence if the fraction of energy in the HEC $hecF > 0.8$ and the number of cells that contain 90% of all energy $n90 <= 5$, the jet tagged as bad. If $hecF > 0.5$ and $Q > 0.5$ the jet is also tagged as bad.

Timing information and shower shape can be used to tag bad jets caused by cosmic rays. If the jet arrives more than 25 ns away from the closest bunch crossing, it is tagged as bad. Note here that all through the 2010 data taking the time between bunch crossings was larger than the design 25 ns. If the fraction of energy of a central jet ($|\eta| < 2$) in one layer of the calorimeter is high (> 0.99), it is probably an atmospheric muon from a cosmic ray interaction. Likewise, if the fraction of energy in the EM calorimeter is less than 0.05, it is likely that the particle traveled from the top of the detector downwards rather than from the interaction point.

When any jet in an event does not meet either of the criteria above the entire event is discarded. A jet can also be tagged as ‘ugly’ when they are found in parts of the detector where measurement of the jet energy is known to be inaccurate, which means the transition region between barrel and endcap, and problematic cells identified during running. These jets do not cause the event to be discarded, but they are not included in the analysis as jets.

3.2 Electrons

In the standard reconstruction of electrons, a seed electromagnetic tower with transverse energy above 3 GeV is taken from the electromagnetic calorimeter. A matching track is searched for among all reconstructed inner detector tracks, which do not belong to a photon-conversion pair. The track after extrapolation to the calorimeter is required to match the cluster within a broad window of $\Delta\eta \times \Delta\phi = 0.05 \times 0.1$, and have a momentum compatible with the cluster energy.

If a match is found, the electron candidate undergoes further quality checks taking into account the shape of the electromagnetic shower, the quality of the inner detector track and the TRT information to discriminate jets from the signal electrons. Three categories of electrons are specified in ATLAS, known as *preselected*, *medium* and *tight*. The exact definition of these types can be found e.g. in [34].

In this analysis medium electrons are used with an isolation requirement. The transverse isolation energy in a cone of $\Delta R < 0.2$ around the electron is required to be smaller than 10 GeV. Electrons in the crack region $1.37 < |\eta| < 1.52$ are not yet well reconstructed in ATLAS. Events with an electron reconstructed in this region are therefore rejected.

As jets and electrons are both objects reconstructed from the calorimeters an overlap removal procedure is defined. Jets reconstructed within a cone $\Delta R = 0.2$ of an identified electron are discarded from the jet list, to prevent double-counting the same object as both a jet and an electron.

For the analysis on 2010 data a relative isolation was required, with a maximum of 15% of the electron p_T allowed in a cone of size 0.2. Furthermore to improve the quality of the tracks used in the electrons, a hit in the b-layer of the Pixel detector is required.

During data taking in 2010 an issue was discovered where the optical link (OTx) that connects the front end electronics of the electromagnetic calorimeter to the ATLAS readout system was malfunctioning. Any electron candidates that pass through a region of the calorimeter with a dead OTx was discarded.

3.3 Muons

Muon are reconstructed using information from both the Muon Spectrometer (MS) and the Inner Detector (ID). The direct approach is to reconstruct *standalone* muons by finding tracks in the MS and then extrapolating these to the beam line. *Combined* muons are found by matching standalone muons to nearby ID tracks and then combining the measurements from the two systems. The reverse, extrapolating inner detector tracks to the spectrometer detectors and searching for nearby hits, results in *tagged* muons. Furthermore *calorimeter* tagging algorithms have been developed to tag inner detector tracks using the presence of a minimum ionizing signal in calorimeter cells, to reconstruct muons in regions uncovered by the muon spectrometer, notably $|\eta| \sim 0$ region where services leave the detector and the barrel/endcap transition region around $|\eta| \sim 1.2$.

The standalone algorithm first builds track segments in each of the three muon stations by performing a straight line fit. Then all possible combinations of at least two segments are linked to form tracks, that are refitted from the drift time measurements of the hits belonging to the segments. This global fit returns a χ^2 value that is used for selection. The resulting standalone tracks are extrapolated to the beam line, correcting for the energy loss in the material in front of the muon spectrometer. The momentum as measured in the muon spectrometer is determined from the bending radius of the muon track by the magnetic field.

Combining of the inner detector and the standalone muon spectrometer tracks is done by the `Staco` algorithm [31] in this analysis. The match chi-square, defined as the difference between outer and inner track vectors weighted by their combined covariance matrix provides an important measure of the quality of this match and is used to decide which pairs are retained as combined muons. On top of a better estimation of the track parameters, the combination of the two detectors also results into a higher rejection of background muons, mainly coming from decays in flight of pions and kaons.

Analysis of 2008 MC simulation

In this analysis a reasonable quality of combined muons was guaranteed with a loose requirement that the tracks should match with $\chi^2 < 100$ and the inner detector track used for the combination be the best match to the track in the muon spectrometer, if multiple inner detector tracks fit. Just as for electrons the muons are required to be isolated, by demanding that the total calorimeter energy deposited in a cone of $\Delta R < 0.3$ around the muon be less than 10 GeV. Muons found to overlap with a jet within $\Delta R < 0.4$ are removed from the collection.

Analysis of 2010 data

The selection cuts for good muons have changed for the analysis of the first ATLAS data. The new cuts are listed in Table 3.1. The goodness-of-fit of the combined track of matched ID and MS hits is required to be better than $\chi^2_{match} < 150$. The distance to the primary vertex is measured in the z direction, along the beamline. The cut on p_T imbalance prevents ID and MS tracks with very different p_T being combined to form a muon, and the isolation cut prevents muons from e.g. b-decay entering the analysis.

In the bottom half of Table 3.1 the requirements of the inner detector track part of the combined muon are listed, demanding a minimum number of Pixel, SCT and TRT hits. There is also a limit on the number of outliers (hits associated with the track, but not used in determining the track parameters since their contribution to the fit χ^2 is too large) relative to the total number of TRT tracks. This requirement is split over different regions in pseudorapidity: in the central part of the detector the number of TRT hits is required to be larger than five, and the relative number of outliers is limited. In the forward region, there is no requirement on the number of TRT hits, but if it is larger than five, the relative number of outliers is limited.

3.4 Missing Transverse Energy

The missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is a very important object in SUSY/mSUGRA analyses, as the presence of a massive weakly interacting particle in SUSY events is the cause of large $E_{\mathrm{T}}^{\mathrm{miss}}$. This can be used to distinguish these events from the SM background.

The measurement of missing transverse energy requires very detailed calibration and measurement of all the sub-detector components of ATLAS. The non-interacting particles, such as neutrinos in the SM and the Lightest Supersymmetric Particle (LSP) for SUSY, carry away energy. To measure this missing energy, all energy deposits in the detector must be measured and calibrated correctly. Since at the LHC the boost of the produced particles along the beam-axis is unknown, one can only speak of missing energy measurement in the transverse plane, or $E_{\mathrm{T}}^{\mathrm{miss}}$.

Analysis of 2008 MC simulation

The $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm first calculates the sum in the two transverse directions, x and y as follows:

$$E_{x,y}^{\mathrm{Final}} = E_{x,y}^{\mathrm{Calo}} + E_{x,y}^{\mathrm{MuonBoy}} + E_{x,y}^{\mathrm{Cryo}}, \tag{3.5}$$

Description	Cut
χ^2_{match} (combined muons only)	$\chi^2_{match} < 150$
Distance to primary vertex	$\|d_z\| < 10\,\text{mm}$
Impact parameter significance	$d/\sigma_d < 5$
Pseudorapidity	$\|\eta\| < 2.4$
p_T imbalance if $p_T^{MS} < 50\,\text{GeV}$	$(p_T^{MS} - p_T^{ID})/p_T^{ID} > -0.4$
Isolation, cone size 0.2	ptcone20$< 1.8\,\text{GeV}$
Distance to closest jet	$\Delta R_{jet} > 0.4$
ID track requirements for muons	
Nr. of SCT hits	$N_{SCT} \geq 6$
Nr. of Pixel hist	$N_{Pix} \geq 1$
TRT outliers $\|\eta\| \leq 1.9$	$N_{h+o} > 5$ AND $N_o/N_{h+o} < 0.9$
TRT outliers $\|\eta\| > 1.9$	$N_{h+o} < 6$ OR $N_o/N_{h+o} < 0.9$

Table 3.1: Muon object definition. The variable d denotes the track impact parameter, σ_d the uncertainty on the impact parameter. The isolation in a cone size of 0.2 is called ptcone20. N_o is the number of outliers, N_{h+o} the number of hits plus the number of outliers.

where

$$E_{x,y}^{\text{Calo}} = -\sum_{\text{TopoCells}} E_{x,y}, \tag{3.6}$$

$$E_{x,y}^{\text{MuonBoy}} = -\sum_{\text{muons}} E_{x,y} \tag{3.7}$$

$$E_{x,y}^{\text{Cryo}} = -\sum_{\text{jets}} w^{\text{Cryo}} \sqrt{E_{x,y}^{\text{EM3}} \times E_{x,y}^{\text{HAD}}}. \tag{3.8}$$

To suppress noise, only calorimeter cells that are associated to a TopoCluster contribute to $E_{x,y}^{\text{Calo}}$. For the muon term, the momenta as measured in the muon spectrometer by the combined algorithm are used to prevent double counting, since the calorimeter term already accounts for the muon energy deposits in the calorimeter. All the combined muons in the $|\eta| < 2.5$ are summed, while for the region uncovered by the inner detector $2.5 < |\eta| < 2.7$ the standalone muon spectrometer momenta are used. The last term of Equation 3.5 accounts for the energy lost in the dead material of the cryostat between the electromagnetic and hadronic calorimeters. The last layer of the electromagnetic LAr calorimeter, $E_{x,y}^{\text{EM3}}$, is compared to the first layer of the hadronic calorimeter, $E_{x,y}^{\text{HAD}}$, and for each jet a calibration weight, w^{Cryo}, is applied.

The $E_{x,y}^{\text{Calo}}$ calculation is further refined by calibrating each contribution according to the reconstructed object it is assigned to. The assignment adheres to the following order: electrons, photons, muons, hadronicaly decaying taus, b-jets and finally light jets. Thus $E_{x,y}^{\text{Calo}}$ ($E_{x,y}^{\text{Final}}$) is replaced by the refined $E_{x,y}^{\text{RefCalo}}$ ($E_{x,y}^{\text{RefFinal}}$) defined as:

$$\begin{aligned} E_{x,y}^{\text{RefCalo}} = \; & E_{x,y}^{\text{RefEle}} + E_{x,y}^{\text{RefGamma}} + E_{x,y}^{\text{RefTau}} + \\ & E_{x,y}^{\text{RefJet}} + E_{x,y}^{\text{RefMuon}} + E_{x,y}^{\text{CellOut}}. \end{aligned} \tag{3.9}$$

As before each term in equation 3.9 is calculated as the negative sum of calibrated cells inside the specific object. All TopoClusters calorimeter cells without an object assignment are collected in the $E_{x,y}^{\mathrm{CellOut}}$ term. In the $E_{x,y}^{\mathrm{RefMuon}}$ term only non-isolated muons are taken into account, as the isolated ones are already in the $E_{x,y}^{\mathrm{MuonBoy}}$ term.

Finally the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is calculated by taking the length of the $E_{x,y}^{\mathrm{RefFinal}}$ vector:

$$E_{\mathrm{T}}^{\mathrm{miss}} = E_{x,y}^{\mathrm{RefFinal}} = \sqrt{(E_x^{\mathrm{RefFinal}})^2 + (E_y^{\mathrm{RefFinal}})^2} \tag{3.10}$$

Analysis of 2010 data

The calorimeter component $E_{\mathrm{T}}^{\mathrm{Calo}}$ definition for the first data analysis is somewhat simplified:

$$E_{x,y}^{\mathrm{SimCalo}} \quad = E_{x,y}^{\mathrm{RefEle}} + E_{x,y}^{\mathrm{RefJet}} + E_{x,y}^{\mathrm{RefMuon}} + E_{x,y}^{\mathrm{CellOut}}. \tag{3.11}$$

Note that the calorimeter cells are calibrated using only Monte Carlo information to account for the difference between hadronic and electromagnetic objects.

In the reconstruction software used in the 2010 analysis the muon term used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ determination was miscalculated, causing large high $E_{\mathrm{T}}^{\mathrm{miss}}$ tails. This has been corrected by removing the muon term from the missing energy and adding by hand the contribution of each reconstructed muon.

3.5 Transverse mass

Another observable that is important for this analysis is *transverse mass*, M_{T}, between an identified lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector, which is defined as:

$$M_{\mathrm{T}} = \sqrt{2(p_T^{lep} E_{\mathrm{T}}^{\mathrm{miss}} - \overrightarrow{p}_T^{lep} \cdot \overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}})}, \tag{3.12}$$

where p_T^{lep} is the transverse momentum of the lepton.

For events that contain a single leptonic W-boson decay, such as semileptonic $t\bar{t}$ and W+jets events, the M_{T} distribution displays a characteristic cut-off around the W mass, $m_W \sim 80\,\mathrm{GeV}$. A good approximation is that the neutrino from the W is the sole particle responsible for the missing transverse energy. When combined with the lepton from the W it results into the transverse W mass, which is at maximum equal to m_W. It is assumed here that the masses of neutrino and lepton are negligible in comparison to their momenta, which is correct for most W-boson decays. For one-lepton SUSY events where $E_{\mathrm{T}}^{\mathrm{miss}}$ is a superposition of the two LSPs and possibly neutrinos from decay of SM particles in the chain, the M_{T} distribution shows no characteristic cut-offs, but is purely defined by the kinematics of the specific supersymmetric model and event selection criteria.

3.6 Three jet mass

The hadronic decay of a top quark results in three hard jets. If one would reconstruct these three jets correctly, one expects to find the top mass peak by calculating the invariant mass of the three jet system. This invariant three jet mass, M_{jjj}, will be used intensively in this analysis, and is defined as:

$$M_{\mathrm{jjj}}^2 = (\sum_{i=1}^{3} p_{max \sum \vec{p}_T}^{jet,i})^2, \tag{3.13}$$

where $p^{jet,i}_{max\sum\vec{p}_T}$ is the four-momentum of the jet i. The $max\sum\vec{p}_T$ subscript denotes that the top quark candidates are selected as the three-jet sum combination with the vector sum of transverse momentum.

In the absence of reliable b-tagging, as was envisioned for the early days of LHC data taking, there is an ambiguity in choosing the correct three-jet combination among the reconstructed jets of a $t\bar{t}$ pair. The combination with the highest transverse momentum chooses the correct pairing in approximately 25% of all events [35], considerably better than random. The ambiguity arises from the fact that semileptonic $t\bar{t}$ production has four partons carrying color charge in the final state and the two top quarks in the initial state, which can all radiate extra partons. The large number of reconstructed jets results in more possibilities to choose a three-jet combination, not necessarily the correct one. On top of that some partons are not reconstructed correctly into a jet or not with the right momentum or not reconstructed at all if they fall outside the acceptance. Another possibility is that all the partons are reconstructed correctly, but if for example the hadronic side bottom quark radiates a hard gluon, its four-momentum will not represent the original top decay daughter, so one would need to take into account the fourth jet of the radiated gluon to reconstruct the top quark mass. The last possibility is that two partons (from a W-boson decay), will be merged into one jet if the boost of the initial particle (W) is very large, pushing the two partons (quarks) close to each other in $\eta - \phi$ space.

3.7 Effective Mass

The effective mass, M_{eff}, is not used as an observable in this analysis but plays a part in the event selection. The definition is:

$$M_{\mathrm{eff}} = \sum_{i=1}^{4} p_T^{jet,i} + \sum_{i=1} p_T^{lep,i} + E_{\mathrm{T}}^{\mathrm{miss}}, \tag{3.14}$$

where the sums run respectively over the four highest p_T jets within $|\eta| < 2.5$ and over all the identified leptons. This variable has the interesting property that for SUSY events the M_{eff} distribution peaks at a value which is strongly correlated with the mass of the pair of SUSY particles produced in the initial proton-proton interaction, which can be used to quantify the SUSY mass-scale.

Chapter 4

Performance of the ATLAS detector on first data

Since commissioning of the ATLAS detector started in 2007, the ATLAS collaboration has made full use of muons from cosmic ray interactions, single beam passes, beam splashes on collimators near the detector and collisions at a wide range of center of mass energies, from 900 GeV to the current maximum energy of 7 TeV. This section shows the performance of the detector, as of fall 2010, after all this information was used to align and calibrate the many components of ATLAS (fall reprocessing).

Table 4.1 quotes the number of channel and overall operational fraction of all the subsystems of the ATLAS detector in 2010. With all channels uptime combined, ATLAS recorded 93.6% of the integrated luminosity delivered by the LHC. More details on the performance of the ATLAS detector can be found on the ATLAS public wiki [36].

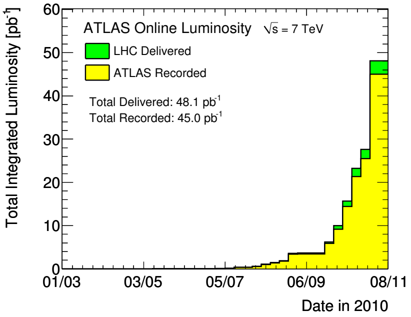


Figure 4.1: The total integrated luminosity delivered to and recorded by the ATLAS detector in 2010.

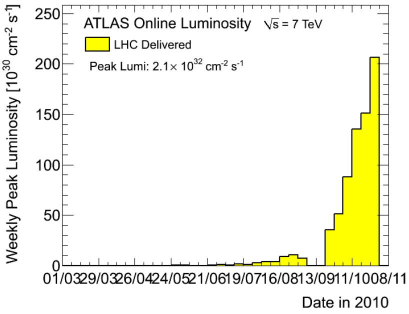


Figure 4.2: The peak luminosity of the LHC per week in 2010.

Detector system	Nr of channels	AOF
pixels	80 M	97.2%
SCT Silicon Strips	6.3 M	99.2%
TRT Transition Radiation Tracker	350 k	97.5%
LAr EM Calorimeter	170 k	99.9%
Tile calorimeter	9800	98.8%
Hadronic endcap LAr calorimeter	5600	99.8%
Forward LAr calorimeter	3500	99.9%
LVL1 Calo trigger	7160	99.9%
LVL1 Muon RPC trigger	370 k	99.5%
LVL1 Muon TGC trigger	320 k	100%
MDT Muon Drift Tubes	350 k	99.8%
CSC Cathode Strip Chambers	31 k	98.5%
RPC Barrel Muon Chambers	370 k	97.0%
TGC Endcap Muon Chambers	320 k	99.1%

Table 4.1: Approximate operational fraction (AOF) of the ATLAS detector systems over the entire 2010 data taking period.

4.1 Starting the LHC

During the first start-up phase of the LHC in september 2008, a faulty electrical connection caused considerable damage to several magnets and other components of the accelerator. After a year of repairs, the second attempt in november 2009 was more successful. Initially the beams were circulated and collided at injection energy, 450 GeV per beam. In december, the energy was ramped up to a center of mass energy of 2.36 TeV. In 2010 the center of mass energy was ramped up to 7 TeV. The luminosity delivered to and recorded by ATLAS is shown in Figure 4.1 as a function of the date in 2010. The analysis of chapter 7 of this thesis is performed over all data recorded in this first year of routine running at high energy. Extra requirements on data quality reduce the available luminosity for analysis to 34.6 pb^{-1}. The peak luminosity per week of the LHC during the 2010 run is shown in Figure 4.2. The peak luminosity over the whole running period was $2.1 \times 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

4.2 Performance of the Inner Detector

The two main responsibilities of the ID are vertex location and track and momentum measurement. Figure 4.3 shows the transverse position of the reconstructed primary vertex, which shows the slightly elongated shape of the beamspot in the y-direction. From this information the vertexing resolution has been determined to be of 30 µm in the transverse plane [37], compared to a design resolution of 18 µm. The resolution along the z-direction is 50 µm[37]. In order to get a good momentum resolution, the alignment of the ID is an important parameter. Figure 4.5 shows the residuals of points along the track, i.e. the distance from a hit in space to the track which is a result of a pattern-recognition and track fitting algorithm. The alignment was improved by reducing these residuals in an iterative alignment procedure. Figure 4.5 shows the residuals before and after the fall reprocessing.

The plots in Figure 4.4 show the reconstructed J/ψ mass using the muon momenta from the

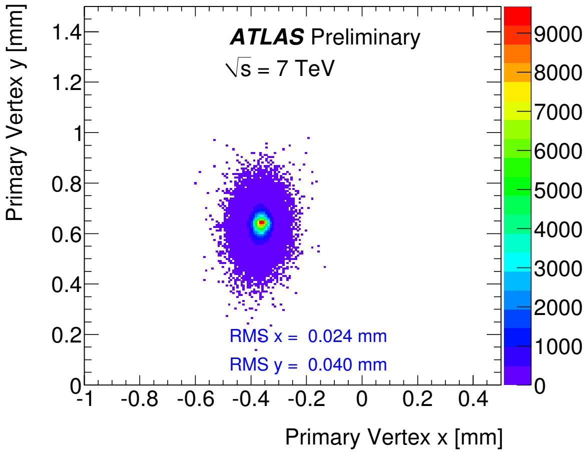


Figure 4.3: Distribution of the measured primary vertex position [37] for the first 6 nb^{-1} taken in spring 2010.

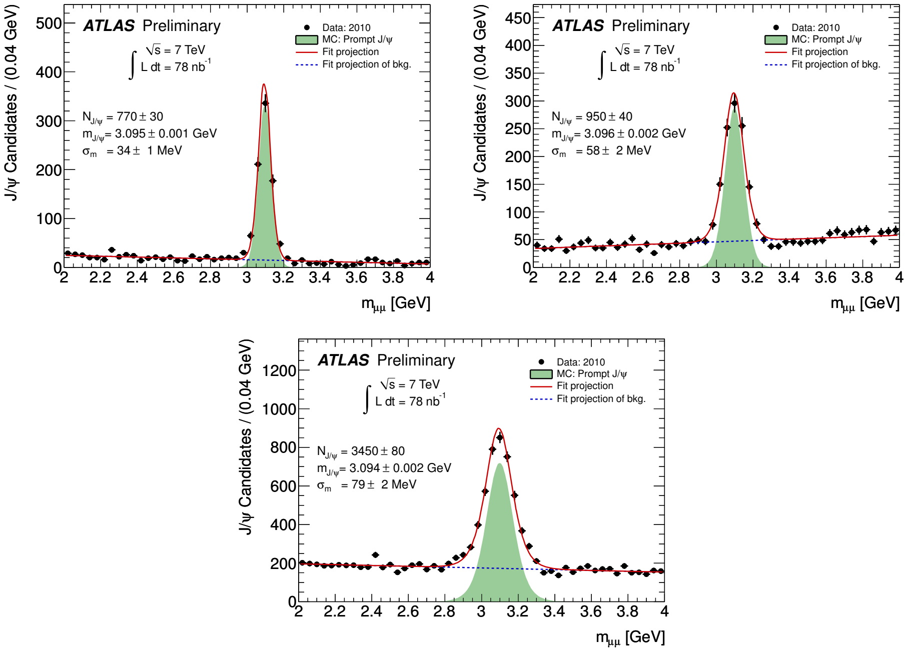
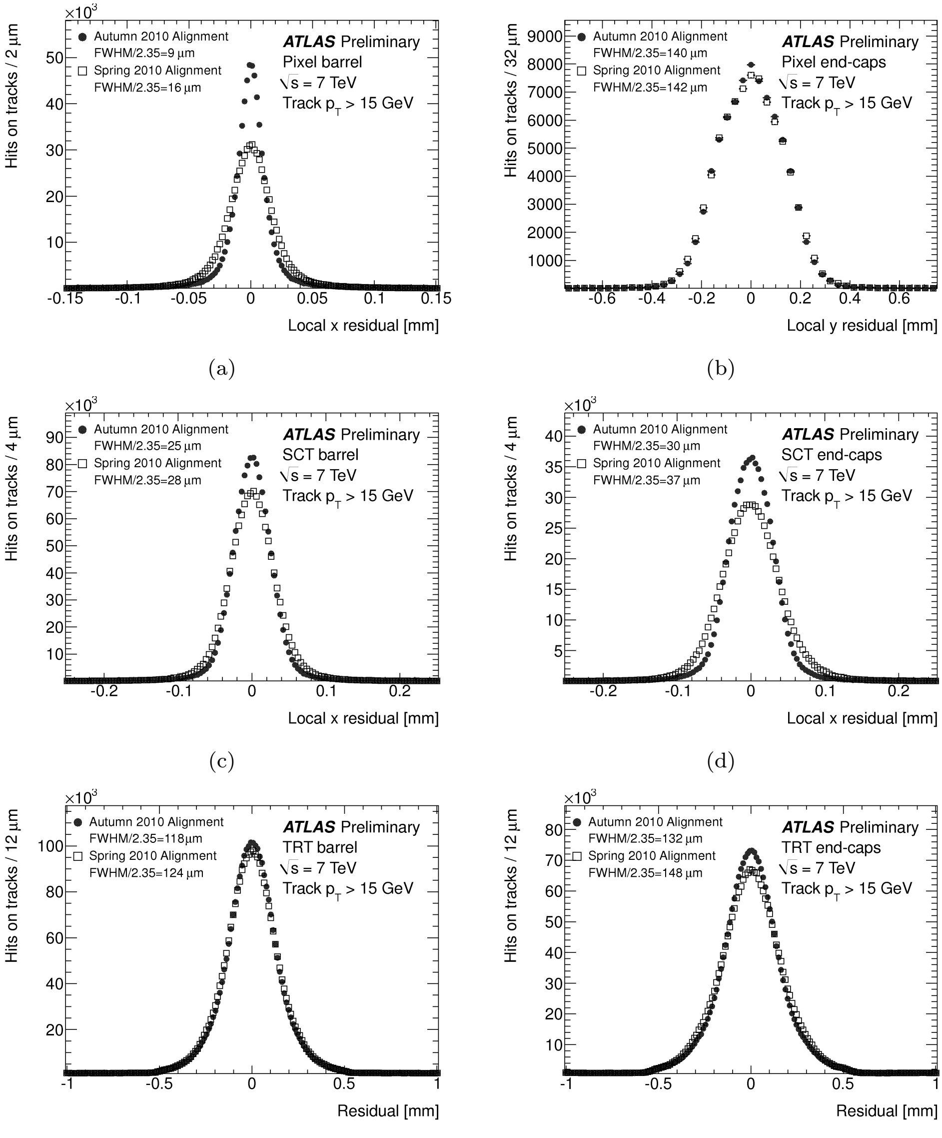


Figure 4.4: The reconstructed mass of the J/ψ with the muon momenta as measured by the ID. The top left plot shows the result with both muons measured in the barrel. The top right plot has one muon in the endcap. The bottom plot has both muons measured by the endcap. Figures taken from [38]. These plots were made before the fall reprocessing.

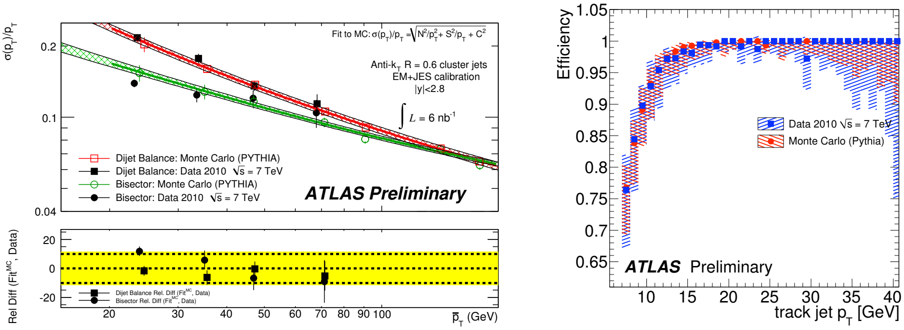


(a) (b)

(c) (d)

(e) (f)

Figure 4.5: These figures show the residuals of the hits in (a) pixel barrel, (b) pixel endcap, (c) SCT barrel, (d) SCT endcap, (e) TRT barrel, (f) TRT endcap. The residuals are defined as the distance from the measured hit location to the expected hit location from the track fit (in the x-direction for the pixel and SCT, and along the ϕ direction for the TRT). The closed circles are the residuals after the fall reprocessing, compared to the open squares from before. Figures taken from ATLAS conference note [39].



(a) (b)

Figure 4.6: Figure (a) shows the jet energy resolution measured with two different methods, compared to MC and fitted with equation 4.1. The bottom plot shows the relative difference between data and MC with a 10% uncertainty (yellow band). Figure (b) shows the jet selection efficiency with respect to track jets as a function of the track jet p_T. These plots were made before the fall reprocessing, and were taken from the ATLAS conference note [40].

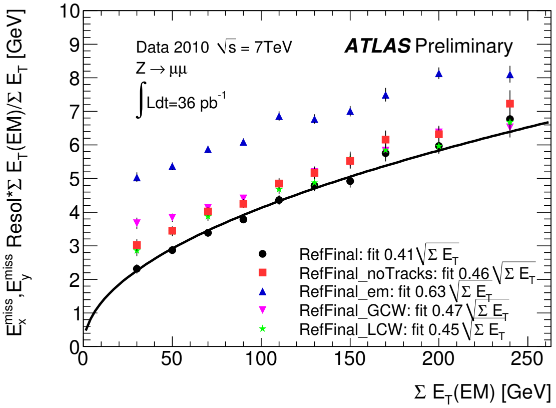
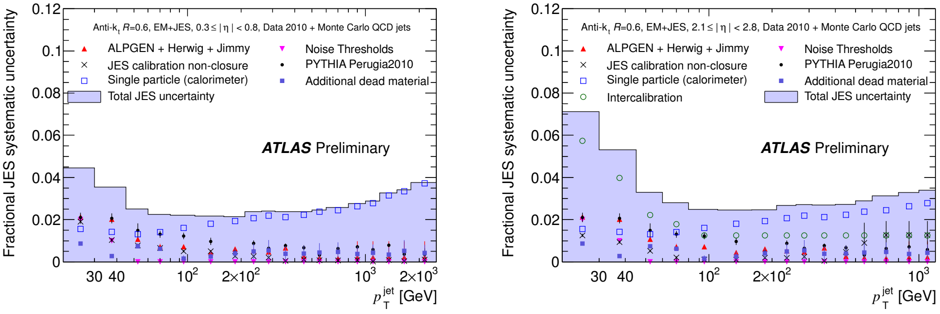


Figure 4.7: The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. Figure taken from the ATLAS conference note [41]. The quoted fit result quotes the fitted coefficient a of the resolution function $\sigma_{E_{\mathrm{T}}^{\mathrm{miss}}} = a\sqrt{E_{\mathrm{T}}^{\mathrm{miss}}}$. The resolution is shown for different calibration schemes. In the analysis of chapter 7 the EM calibration is used. This plot was made after the fall reprocessing.

ID. The top left plot shows that the width of the distribution and its central value for muons in the barrel is consistent with the expectation from MC, indicating that the momentum scale and resolution are as expected. The top right and bottom plots, where one, respectively both, muons are found in the endcap show that the momentum resolution for the endcap is not yet as expected.

For the Pixel detector the remaining uncertainty due to misalignment is 9 µm for the barrel and 15 µm for the endcap. For the SCT the remaining misalignment uncertainty is and 25 µm for the barrel and 30 µm for the endcap respectively. The alignment uncertainty of the TRT endcap is 118 and 132 µm for barrel and endcap [39].

4.3 Performance of the Calorimeter



(a) (b)

Figure 4.8: Systematic uncertainty on the Jet Energy Scale for (a) the barrel and (b) the endcap. Taken from the ATLAS JES conference note [42].

For the analysis presented in this thesis there are three important performance parameters of the calorimeter, namely jet energy resolution, jet energy scale (JES) and $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. The jet and electron energy resolution is shown in Figure 4.6(a), comparing the expectation from MC with the result for data for two different methods. The first uses momentum balance in di-jet events to determine the jet energy scale. The second (called the bi-sector method) is an improvement on this method which takes into account variations perpendicular to the jet axis as well [40]. The two methods differ slightly because the bi-sector method is unaffected by imbalances at particle level, and in general provides a better estimate for the jet energy resolution.

The same figure shows the expected jet energy resolution from MC simulation as a function of p_T, which agrees with the measured jet energy resolution within 18% for jets with p_T between 5 and 40 GeV [40]. The two lines in the plot show the fit to the result of both methods of the jet energy resolution function:

$$\frac{\sigma_{p_T}}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus C, \tag{4.1}$$

with a band around the line indicating the uncertainty of the fit. In this function N, S and C denote the noise, stochastic and constant term respectively.

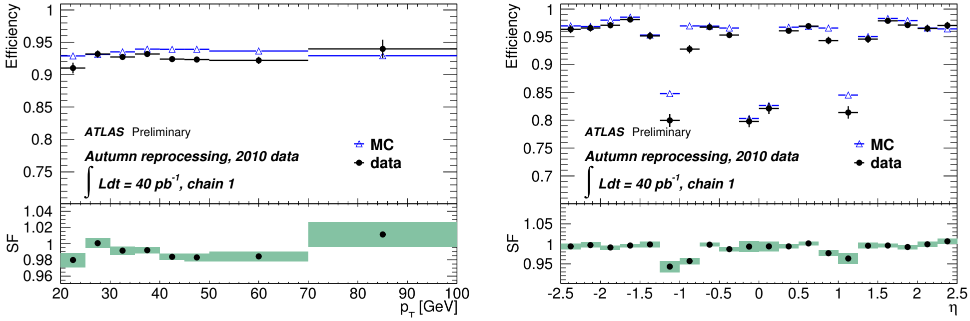
Figure 4.6(b) shows the reconstruction efficiency of the anti-k_t algorithm with respect to track jets as a function of track jet p_T. Both figures were made before the fall reprocessing.

The missing transverse energy is calculated from three components: total transverse energy in the calorimeter, energy loss in the cryostat and muon energy. The resolution of the calorimeter component of the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss,calo}}$) is shown in Figure 4.7. This figure shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution for different calibration schemes. For the analysis of chapter 7 EM calibration was used.

The systematic uncertainty on the JES is determined by the uncertainty of the single hadron interaction with the calorimeter material and of the MC simulation of the jet response

[42]. Figure 4.8(a) shows the systematic uncertainty on the JES as a function of jet energy in the $0.3 < |\eta| < 0.8$ region of the detector [1]. The different sources of uncertainty that were considered are shown. The single particle response is the biggest source of uncertainty for high p_T jets. The results for this region in η, for which most data from test-beams is available, are extrapolated to other η regions in the detector. Figure 4.8(b) shows the effect on the JES uncertainty in the endcap, with the η inter-calibration becoming the dominant source of systematic error at low jet p_T, while at high p_T the single particle response remains the largest source of uncertainty.

4.4 Performance of the Muon Spectrometer



(a) (b)

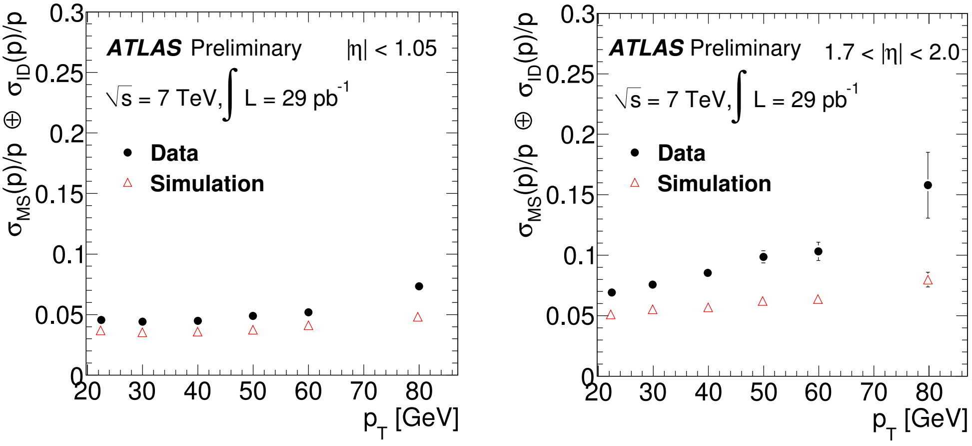
Figure 4.9: Muon reconstruction efficiency as a function of p_T (a) and η (b) after the fall reprocessing [43].

Like the Inner Detector, the Muon Spectrometers' main task is to measure position and momentum of particles. Besides that, it serves as particle ID for muons. The efficiency of finding and reconstructing muons was measured using a so-called 'Tag and Probe' method at the Z resonance. This method selects two oppositely charged tracks with an invariant mass around the Z mass. If one of these muon (the 'tag') forms a combined muon, the other ID track can be used as a probe to determine the MS efficiency of finding the second muon. The result of this method is shown after compensating for background to the Z signal in Figure 4.9 as a function of muon p_T and η. Chain 1 in these plots refers to the specific muon reconstruction algorithm that was used to make this plot, which is the same algorithm as used in the rest of this thesis. As can be seen from Figure 4.9, the muon reconstruction efficiency in data is slightly lower than expected from MC. The dips in the muon reconstructed efficiency as a function of η seen in Figure 4.9(b) are caused by the ATLAS support structure at $|\eta| = 1.2$ and a feedthrough of electronic cables and other services at $\eta = 0$.

The momentum resolution of the MS is shown in Figure 4.10, for both the endcap and the

---

[1] Although this figure shows the uncertainty on anti-k$_t$ jets with a size R=0.6, the uncertainty on the jets used in this analysis (with R=0.4) is very similar.



(a)

(b)

Figure 4.10: The momentum resolution of the Muon Spectrometer after the fall reprocessing plus the ID as a function of muon p_T for the barrel (a) and endcap (b). Figures taken from [44]

.

barrel. These figures show the combined muon p_T resolution:

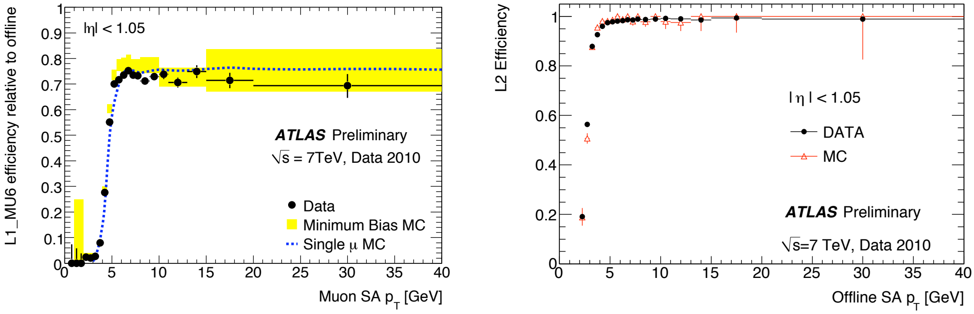
$$\frac{\Delta p}{p} = \sqrt{\frac{\Delta p_{ID}}{p_{ID}}^2 + \frac{\Delta p_{MS}}{p_{MS}}^2}, \tag{4.2}$$

where p_{ID} is the momentum measured by the Inner Detector, and p_{MS} is the momentum measured by the Muon Spectrometer.

The result for MC is plotted in the same figures, showing that the momentum reconstruction works slightly better on MC. The relatively poor resolution in the endcap is caused by the low acceptance of cosmic muons in the endcap region [44], causing the alignment of the endcap modules to be less than optimal.

4.5 Performance of the Trigger

This section will not discuss the performance of the whole trigger system, but focus on the muon triggers which were used in this analysis. To study the muon trigger, samples were selected using two unrelated triggers, the L1 calorimeter and minimum bias trigger, which takes information either from dedicated Minimum Bias Trigger Scintillators installed in the forward region and from central tracking detectors which trigger on randomly selected filled bunches. Muons that are reconstructed offline are matched to a L1 muon Region of Interest (RoI). The same technique was used for the L2 trigger and EF. Figure 4.11(a) shows the efficiency as a function of offline muon p_T of the L1 muon trigger with a p_T threshold of 6 GeV, relative to the offline reconstructed muons in an unbiased sample. The same figure shows the results for minimum bias and single-muon MC sample. The L1 muon trigger performance is well described by MC simulation.



(a) (b)

Figure 4.11: Figure (a) shows the L1 efficiency for muon trigger with 6 GeV threshold (MU6) with respect to offline muons, as a function for offline MS-only (stand-alone, SA) muon p_T. Figure taken from conference note [45]. Figure (b) shows the efficiency with respect to the L1 for the L2 MSonly muon triggers with a p_T threshold of 4 GeV, as a function for offline MS-only (stand-alone, SA) muon p_T, for the MS barrel (a) and endcap (b). Figure taken from conference note [45].

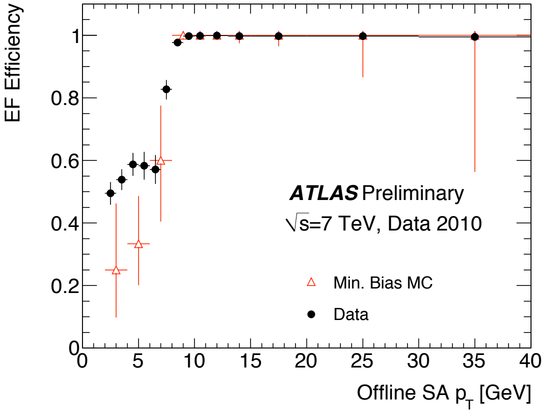


Figure 4.12: Efficiency for the EF muon trigger with a threshold of 10 GeV with respect to L2, as a function for offline MS-only (stand-alone, SA) muon p_T. Figure taken from conference note [45].

Figure 4.11(b) shows the efficiency as a function of muon p_T for the L2 muon trigger with a threshold of 4 GeV, relative to offline reconstructed muons matched to a L1 RoI. This L2 trigger takes information not only from the MS, but also from the ID, making a better estimation of the muon momentum possible. Figure 4.12 shows the efficiency of the standalone muon EF with 10 GeV p_T threshold with respect to the L2. Again MC simulation describes the measured performance of the trigger systems rather well.

Chapter 5

Conditional product PDFs

5.1 Introduction

The collection of measurements of all collision products that are caused by one or more pp collisions in one bunch crossing in the center of ATLAS is called an event. Each event is a discrete occurrence in time and has one x or more $\vec{x}$ measured observables associated with it. A statistical analysis of many of these events deals with the distribution of measured observables. We are interested in a search for events of a certain origin, which we call signal events, but our search will be affected by the presence of events from a different origin which we call background events. Probability density functions (PDFs) will be used to describe a model of the behavior of signal and background distributions, and using a maximum likelihood estimation method the parameters of this model can be determined from data. This section will deal with some of the mathematics involved when working with PDFs. These methods will then be applied in chapters 6 and 7 to perform an actual search for new physics.

A PDF $F(\vec{x};\vec{p})$ gives the probability density of a distribution of observables $\vec{x}$, where the vector $\vec{p}$ contains the parameters of the model describing this distribution. A probability density function obeys two rules: it must be normalized to unity and it must be positive definite for all possible values of $\vec{x}$ and $\vec{p}$.

As an example we take a Gaussian distribution in observable x. The parameters of the model are the mean (μ) and standard deviation (σ). The PDF $G(x;\mu,\sigma)$ following the Gaussian function can be written as:

$$G(x;\mu,\sigma) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\displaystyle\int_{x_{min}}^{x_{max}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx}, \tag{5.1}$$

where $x_{min}(x_{max})$ are the lower (upper) limit of the observable x.

There are many advantages to using PDFs. First of all they are a prerequisite for the use of an (unbinned) maximum likelihood estimation technique. For a given set of measurements x_i the likelihood is defined as a product of the PDF value for each measured value of x:

$$L(\vec{p}) = \prod_i F(\vec{x}_i;\vec{p}). \tag{5.2}$$

The likelihood function thus gives the probability of finding the set of data points x_i for a given set of parameters. Finding the maximum of the likelihood means finding the set of parameters for which this measurement is most likely, thus finding the set of parameters that best describe this dataset. For convenience the negative log of the likelihood is often used:

$$-\ln L(\vec{p}) = \sum_i \ln F(\vec{x}_i;\vec{p}), \tag{5.3}$$

as it is computationally easier to minimize this sum, which is equivalent to maximizing the likelihood $L(\vec{p})$.

Another advantage to the use of PDFs is that they can be added together with an intuitive interpretation of fraction coefficients. In order to describe for example a signal distribution $S(x)$ on top of a background $B(x)$ with a fraction α, a sum of PDFs can be used:

$$F(x) = \alpha S(x) + (1-\alpha)B(x), \tag{5.4}$$

defining $\alpha \in [0,1]$ which can be another parameter in the model. Because both the signal and the background PDFs are normalized to one, by construction the sum of these PDFs is also normalized to unity. For shorter notation $F(x)$ will be used from now on to describe a PDF leaving out the parameters wherever this is convenient.

A model can also be extended to multiple observables or dimensions by simply writing a product of lower or one-dimensional PDFs as:

$$H(x,y) = F(x) \cdot G(y), \tag{5.5}$$

where $F(x)$ and $G(y)$ are factorizing PDFs. In other words: the parameters of $F(x)$ and $G(y)$ are independent of y and x respectively.

If, on the other hand, the observables x and y are correlated, the language of PDFs can be used to describe these correlations by making the parameters of one dimension depend on the observable of the other dimension:

$$H(x,y) = F(x;\vec{p}(y)) \cdot G(y) = F(x|y) \cdot G(y), \tag{5.6}$$

using the notation $F(x|y)$ to indicate that this PDF is a *conditional PDF* in the y-dimension. This means $F(x|y)$ describes the distribution in x for a given value of y, but the PDF itself has no information on the y distribution. The PDF $H(x,y)$ is called a *conditional product PDF*. Details on how to implement such a PDF are given in the next sections.

Finally some clarification of used notation: lowercase letters are used $f(x;p)$ to denote an unnormalized PDF, technically a function. Uppercase $F(x;p)$ are reserved for a properly normalized PDF. Where needed the range R over which the PDF is normalized is denoted using a subscript: $F_R(x;p)$.

5.2 Shape of a conditional product PDF

When a conditional product PDF is defined as in equation 5.6, the shape of the PDF will depend on the range in the observable x on which the PDF is normalized. This behavior is explained in this section.

The normalization of a conditional PDF $F(x|y)$ is different from a standard 2-dimensional PDF $H(x,y)$ by the normalization:

$$\int_R F_R(x|y)dx = 1 \tag{5.7}$$

$$\iint_R K_R(x,y)dxdy = 1. \tag{5.8}$$

The conditional PDF integrates to unity over x in range R for each value of y. Thus the PDF itself does not have any predictive power in the y dimension. The y-dependence of this PDF rests solely in the y-dependence of its parameters. This is achieved through the normalization factor $N(y)$ specific to conditional PDFs:

$$N(y) = \frac{1}{\int_R f(x,y)dx} \tag{5.9}$$

$$F(x|y) = N(y)f(x,y) \tag{5.10}$$

with $f(x,y)$ the unnormalized conditional function. This normalization factor of course depends on y.

A conditional PDF describes the shape of the model in only one dimension (x), but needs information on y. To add a second dimension, the conditional PDF $F(x|y)$ can be multiplied by a PDF $G(y)$ to form a conditional product PDF:

$$H(x,y) = F(x|y)G(y) = \frac{f(x,y)g(y)}{\int f(x,y)dx \int g(y)dy}. \tag{5.11}$$

Integrating $H(x,y)$ over x and y (noting the fact that the denominator has a remaining y-dependence) shows that it is properly normalized, i.e. it is a proper PDF. As each component of $H(x,y)$ is normalized separately, the normalization of $H(x,y)$ is defined as a product of two 1D integrals. This has consequences which will now be explored.

An example of a conditional product PDF is shown in the left plot of Figure 5.1. This plot shows a Gaussian PDF in x multiplied by a uniform PDF in y.

$$H(x,y) = \text{Gaussian}(x, \mu(y) = y, \sigma = 3) \cdot \text{Uniform}(y). \tag{5.12}$$

The PDF is normalized on the range $x \in [-10, 10]$, $y \in [-10, 10]$. The mean of the Gaussian is a linear function of y. This gives the PDF a sliding peak visible in Figure 5.1. The normalization factor as defined in equation 5.9 of the conditional component $F(x|y)$ of this PDF is given by

$$N(y) = \frac{1}{\displaystyle\int_{-10}^{10} e^{-\frac{1}{2}\left(\frac{x-y}{3}\right)^2} dx}. \tag{5.13}$$

Since $F(x|y)$ must be normalized to 1 for every y, the shape of the conditional product PDF will change when the normalization range in x changes, as is apparent when looking at the definition of the normalization constant in equation 5.13. The y-dependence of $N(y)$ changes for a different integration range.

As an example, the normalization factors for two different normalization ranges are plotted in Figure 5.2. The solid line shows the normalization when the PDF is normalized over $x \in$

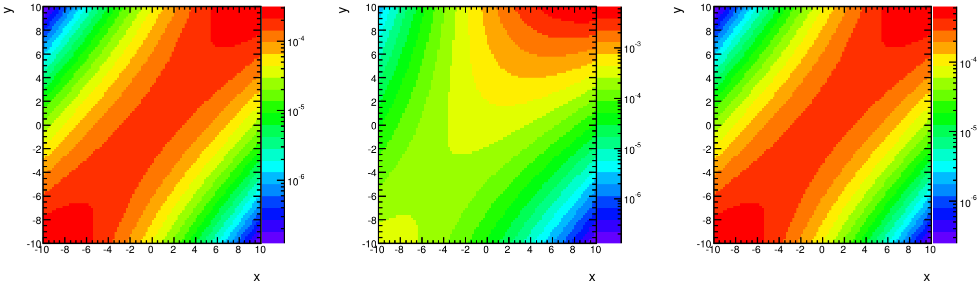


Figure 5.1: This figure shows a uniform PDF in y multiplied by a Gaussian in x whose mean is a linear function of y. On the left, is the original model defined in the range $x \in [-10, 10]$. In the middle is the same model, but now normalized to a subrange $x \in [-10, 0]$. On the right is the same model, normalized the subrange, but with the reference range set to $x \in [-10, 10]$

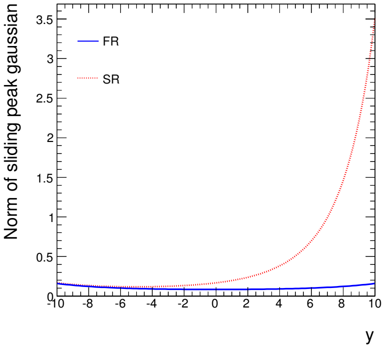


Figure 5.2: The normalization factor for the model shown in Figure 5.1 normalized on $x \in [-10, 0]$ (dotted line) compared to the normalization factor over $x \in [-10, 10]$ (solid line), as a function of y.

$[-10, 10]$, the dashed line shows the normalization constant when normalizing over $x \in [10, 0]$. The difference is clear: for higher values of y, the normalization factor for the range $x \in [-10, 0]$ becomes very large. The reason is that at higher y the peak of the Gaussian model in x lies outside the normalization range. Thus to keep the normalization to unity for all values of y as required for a conditional PDF, the normalization factor must increase as a function of y.

The result of the different normalization factor is shown in the middle plot of Figure 5.1. This shows the same model, a Gaussian with a sliding peak, but now it is normalized to the range $x \in [-10, 0]$. For comparison, the PDF is still plotted in the full range $x \in [-10, 10]$, although technically it is only valid as a PDF in the normalization range. Notice that the shape of the PDF has changed completely. This means that if this model was used to compare e.g. the relative number of events in the range $x \in [-10, 0]$ to the number of events in the range $x \in [0, 10]$ the result would depend on the normalization range that was used.

This is not desired behavior if one wants to define a model on multiple ranges, e.g. when extrapolation of the model from one range to another is required.

Thus a *reference range* must be defined which defines the shape of the PDF, and a separate range which defines the normalization (*normalization range*). The result of this is shown in Figure 5.1 on the right. Here, the PDF is normalized on $x \in [-10, 0]$, but the reference range defining the shape of the model is set to $x \in [-10, 10]$. Comparing the color axis of the first and last plots shows that indeed the overall normalization is different, but the shape of the two distributions is clearly the same.

Implementing a reference range means we no longer see the components of the conditional product PDF of equation 5.11 as separate PDFs. Rather they form one object which has to be normalized as a whole. The next paragraph will discuss how this reference range and normalization is implemented.

5.3 Conditional product PDF with reference range

Requiring that the shape of the PDF is the same regardless of the range over which it is normalized is equivalent to the requirement that the ratio

$$R = \frac{\iint_{R1} F_N(x|y) \cdot G_N(y)dxdy}{\iint_{R2} F_N(x|y) \cdot G_N(y)dxdy} \tag{5.14}$$

for any pair of ranges $R1$ and $R2$ is the same regardless of the normalization range N of F and G. In this section it will be shown that this is not the case when one uses conditional product PDFs. First insert explicitly the normalization of $F(x|y)$ and $G(y)$ over normalization range N:

$$R = \frac{\iint_{R1} \frac{f(x,y)}{\int_N f(x,y)dx} \frac{g(y)}{\int_N g(y)dy} dxdy}{\iint_{R2} \frac{f(x,y)}{\int_N f(x,y)dx} \frac{g(y)}{\int_N g(y)dy} dxdy}. \tag{5.15}$$

The normalization integral of $g(y)$ is of course a constant, so it cancels between numerator and denominator. However, the normalization of $f(x, y)$ still depends on y, which means that the ratio still depends on the normalization range N:

$$R = \frac{\iint_{R1} \frac{f(x,y)}{\int_N f(x,y)dx} g(y)dxdy}{\iint_{R2} \frac{f(x,y)}{\int_N f(x,y)dx} g(y)dxdy}. \tag{5.16}$$

To solve this the definition of a conditional product PDF has to be adapted to fix this unwanted behavior. We substitute the unnormalized function $g(y)$ of the non-conditional component $G_N(y)$ by:

$$g(y) \to g(y) \cdot \frac{\int_N f(x, y)dx}{\int_{ref} f(x, y)dx}, \tag{5.17}$$

where ref denotes the reference range which will define the shape of the conditional product PDF.

The rest of this section will show that by doing this replacement the conditional product PDF now obeys the requirement that the ratio of equation 5.14 is independent of the normalization range N. Performing the substitution of equation 5.17 in the formula for the ratio 5.16 gives:

$$R = \frac{\iint_{R1} \frac{f(x,y)}{\int_N f(x,y)dx} g(y) \frac{\int_N f(x,y)dx}{\int_{ref} f(x,y)dx} dxdy}{\iint_{R2} \frac{f(x,y)}{\int_N f(x,y)dx} g(y) \frac{\int_N f(x,y)dx}{\int_{ref} f(x,y)dx} dxdy}. \tag{5.18}$$

The two integrals over the normalization range cancel, leaving:

$$R = \frac{\iint_{R1} \frac{f(x,y)}{\int_{ref} f(x,y)dx} g(y)dxdy}{\iint_{R2} \frac{f(x,y)}{\int_{ref} f(x,y)dx} g(y)dxdy} = \frac{\iint_{R1} F_{ref}(x|y)g(y)dxdy}{\iint_{R2} F_{ref}(x|y)g(y)dxdy}, \tag{5.19}$$

where F_{ref} is defined as:

$$F_{ref} = \frac{f(x,y)}{\int_{ref} f(x,y)dx}. \tag{5.20}$$

Equation 5.19 obeys the requirement that the ratio is independent of the normalization range N. Hence it has been shown that by implementing the substitution 5.17 the shape of the PDF is now independent of the normalization range.

Now the substitution of equation 5.17 can be inserted in the 2D conditional product PDF of equation 5.11:

$$H_N^{ref}(x,y) = \frac{f(x,y)g(y)}{\int_N f(x,y)dx \int_N g(y)dy} \rightarrow \frac{f(x,y)g(y)\frac{\int_N f(x,y)dx}{\int_{ref} f(x,y)dx}}{\int_N f(x,y)dx \int_N \frac{\int_N f(x,y)dx}{\int_{ref} f(x,y)dx} g(y)dy}. \tag{5.21}$$

The integrals over the normalization range N in denominator and numerator cancel, leaving:

$$H_N^{ref}(x,y) = \frac{F_{ref}(x|y)g(y)}{\iint_N F_{ref}(x|y)g(y)dxdy}. \tag{5.22}$$

This equation shows that the product PDF of two separate components $f(x,y)$ and $g(y)$ is redefined as a single 2-dimensional object $F_{ref}(x|y)g(y)$ that has a fixed shape defined by the reference range. The normalization of this object is defined by an integral in two dimensions. Effectively $f(x,y)$ was replaced by $F_{ref}(x|y)$ and taken inside the integral over y.

We would like to draw attention to the fact that on a subrange in the observable of the conditional PDF (in this case x) the denominator can only be calculated numerically, because the conditional PDF is no longer normalized to unity for each value of y. Only the entire 2D PDF retains its proper normalization.

5.4 Reference range in a composite subrange

Normalizing a PDF on a 2-dimensional range is easiest when the normalization range is a square in the plane spanned by the two observables. If it is not square, it might be possible to compose it from square elements, as is shown in Figure 5.3.

So far, the normalization of a conditional product PDF was defined in terms of the normalization of the 1D components:

$$H_N(x,y) = F_N(x|y)G_N(y). \tag{5.23}$$

One would be tempted to define the normalization over a range $N = A + B$ by separately normalizing F and G over this composite range. However, when the recipe of the previous section is applied, the normalization of $F(x|y)$ and $G(y)$ are tied together as in equation 5.21, and this integral separation cannot be done.

The solution to this is to normalize the conditional product PDF $H(x,y)$ as a single 2-dimensional object instead of normalizing each component separately. So instead of writing

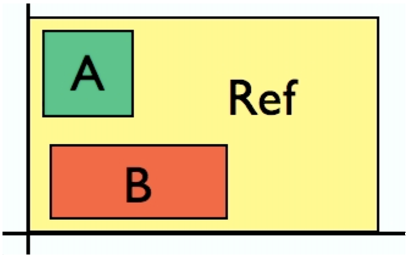


Figure 5.3: Composite ranges within the reference range.

equation 5.23 with $N = A + B$, the PDF in equation 5.21 now normalized over a composite range becomes:

$$H_{A+B}(x,y) = (F_{ref}(x|y)g(y))_{A+B} = \frac{F_{ref}(x|y)g(y)}{\iint_A F_{ref}(x|y)g(y)dy + \iint_B F_{ref}(x|y)g(y)dy}. \tag{5.24}$$

This normalization equation can be extended to an arbitrary number of ranges that make up the composite normalization range:

$$H(x,y)_{\Sigma R} = \frac{F_{ref}(x|y)g(y)}{\Sigma_R \iint_R F_{ref}(x|y)g(y)dy}. \tag{5.25}$$

5.5 Product of a sum of conditional PDFs

First consider take a PDF in x (not conditional for now) which is a sum of two PDFs:

$$F_N(x) = \alpha_N J_N(x) + (1 - \alpha_N)I_N(x), \tag{5.26}$$

noting that such a sum of PDFs is only properly defined if the components J and I are normalized on the same region N in phase space. When the PDF F is extrapolated to a range that is different from N, the coefficient α will change. Therefore the coefficient has an associated range just like the PDF, which is denoted in equation 5.26 with the subscript N. The coefficient can be extrapolated simply like:

$$\alpha_C = \frac{\int_C \alpha_N J_N(x)dx}{\int_C F_N(x)dx}. \tag{5.27}$$

The ranges N and C in this equation can be made composite in a trivial manner. This transformation must be taken into account when dealing with sums of PDFs where one of the PDFs is conditional.

Starting from the 1D PDF given by equation 5.26, dependence on a second observable can be added like this:

$$K_N(x,y) = F_N(x) \cdot G_N(y) = (\alpha_N J_N(x) + (1 - \alpha_N)I_N(x))G_N(y). \tag{5.28}$$

Writing out the normalization explicitly gives:

$$K_N(x,y) = \left(\alpha_N \frac{j(x)}{\int_N j(x)dx} + (1 - \alpha_N)\frac{i(x)}{\int_N i(x)dx}\right) \cdot \frac{g(y)}{\int_N g(y)dx} \tag{5.29}$$

If any of the three components is a conditional PDF, it is best to treat the sum of PDFs as one object. As an example, take $J(x|y)$ as a conditional PDF. F then becomes a conditional PDF as well. Hence the normalization of $K(x, y)$ must be changed according to the recipe of equation 5.25:

$$H_N(x, y) = (\alpha_{ref} J_{ref}(x|y) + (1 - \alpha_{ref}) I_{ref}(x)) \frac{g(y)}{\iint_N F_{ref}(x|y) g(y) dx dy}. \tag{5.30}$$

Note that everywhere the sum PDF remains normalized consistently, but the normalization and the coefficient ranges of F have been changed to the reference range. It is now straightforward to extend this normalization to a composite range, by replacing N in the denominator with a sum over ranges ΣR.

Chapter 6

The Combined Fit method for background estimation to supersymmetry

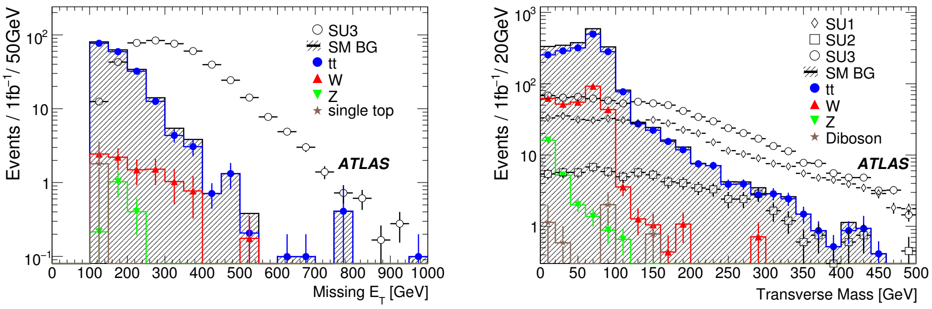
6.1 Introduction

In this chapter we will be looking at a technique to find new physics, in particular supersymmetry, using the ATLAS detector. Before making a statement on new physics signatures that can be found, we have to be sure that the known SM physics at the LHC scale is understood and under control. Existing MC generators, which are incorporated in the ATLAS software framework, can be a helpful tool, but as these have never been tested at such high energies, it is not sufficient to rely only on these generators. Instead, we will be employing a technique to derive as much information from data as possible, using PDFs as described in chapter 5.

The starting point for our strategy is a counting experiment: we want to determine the number of SM events we expect in a certain phase space region, and we compare this number of expected events to the number of events found in the data. To fix this region where we do the experiment, called the signal region, we look at two observables where the SUSY spectrum typically extends to much higher values than the spectrum of SM background. These observables are the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass M_{T}, see Figure 6.1. By cutting at a certain, high value of $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, we can get a good signal over background ratio (S/B) in our signal region.

To determine the background contribution in this region of high $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, we use the fact that this region is 2-dimensional. We can now define an L-shaped control region as in Figure 6.2 in the $E_{\mathrm{T}}^{\mathrm{miss}}$-$M_{\mathrm{T}}$ plane, which ideally is free of SUSY. Thus we can see the full shape of the M_{T} distribution of the data at low $E_{\mathrm{T}}^{\mathrm{miss}}$, and the full $E_{\mathrm{T}}^{\mathrm{miss}}$ shape at low M_{T}. Extrapolating this into the signal region will give us an estimate of the number of SM background events in our signal region.

To model the background we look for a parameterized PDF that is capable of describing the M_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the MC generated samples, yet is very flexible to account for discrepancies between simulated and measured data. We added a third observable, M_{jjj} (see section 3.6), which helps us determining the contribution of the semileptonic $t\bar{t}$ background.



(a) (b)

Figure 6.1: Missing transverse energy (a) and transverse mass (b) for the main backgrounds in one-lepton channel and some ATLAS SUSY benchmark points (SU1, SU2, SU3). Though these distributions were made for $\sqrt{s} = 14$TeV [31], the 10 TeV distributions are quite similar.

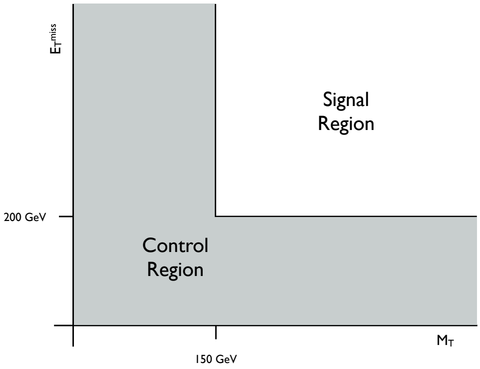


Figure 6.2: The L-shaped control region in the E_T^{miss}/M_T plane.

We keep dependence on MC as small as possible by parametrizing the shapes of these PDFs. We perform an unbinned maximum likelihood fit on the data in the control region to fix these parameters from data.

There are two issues with this naive approach, which we address. Firstly, if SUSY exists there will be a non-negligible amount of SUSY events in our control region. To account for this SUSY contamination we need to model our background *and* our signal in the control region to separate these two contributions and to be able to extrapolate the background model only to the signal region.

As it turns out, it is possible to model a wide range of different mSUGRA models with a simple two-parameter model for low E_T^{miss} and M_T, as will be described in section 6.4.4. By combining our SUSY model with our background model, we can fit the shape parameters and the ratio of signal to background events in our control region at the same time. Once we have determined these parameters, we can extrapolate the background model to the signal region to get an estimate of the expected number of background events in the signal region. Note that although we may fit a SUSY contribution in the control region, our measurement is in the signal region only. We will have to make sure that when no SUSY particles are produced, no SUSY contribution is fitted in the CR introducing a bias in our estimate.

There is a second issue with the naive approach, which is that the shape of the M_{T} ($E_{\mathrm{T}}^{\mathrm{miss}}$) distribution might be different for lower or higher $E_{\mathrm{T}}^{\mathrm{miss}}$ (M_{T}). This turns out to be the case for some backgrounds and we have to take these correlations into account as described in section 6.4.2.

The first section will describe the samples that were generated for the backgrounds and possible mSUGRA signals.

In the second section we will describe the models for the different backgrounds and for the SUSY signal. In this analysis all the aspects of working with PDFs mentioned in chapter 5 are used. A product of three PDFs is used to get a 3-dimensional model of the backgrounds and signal, taking correlations between observables into account using conditional PDFs. Different aspects of a 1D PDF are described by summing PDFs, for instance a Gaussian peak on top of an exponential tail in the M_{T} models. In 3 dimensions, the models of all background contributions and the SUSY signal are combined through addition into one model describing the data.

The control region used in this analysis is non-square, but rather built from two square elements. Together with the signal range that means that the model has to be consistently defined and evaluated in three different ranges. Thus the tools and formalisms developed in the previous chapter will be needed. All the descriptions of the models, and all fitting in this thesis was done using the `RooFit` framework [46], which is part of the `ROOT` framework [47] commonly used in high energy physics.

Finally, a proof of principle will be given where the combined fit method will be applied to a pseudo dataset made by combining background and signal MC. It will be shown that the input yields of the different components can be correctly estimated with the method.

This analysis was developed together with A. Koutsman. Most of the text in this chapter was already published in the ATLAS internal note [48], which is only available for members of the ATLAS collaboration, and appeared as a chapter in Koutsman's thesis [49].

6.2 Signal and Background generation

In the following section we briefly describe our simulation of signal and background samples. The simulated data documented here has been produced by Monte Carlo (MC) simulation inside the official ATLAS software framework `ATHENA` [50]. Then we will describe the criteria for event selection, and the number of events for each background are expected to pass these cuts.

The `ATHENA` framework incorporates different packages for generating events. Simulation of the ATLAS detector was done using the ATLFAST2 [51] simulation package. ATLFAST2 includes state of the art GEANT4 [52] simulation of the inner detector and muon system supplemented by a fast calorimeter simulation. GEANT4 handles the interaction of the particles with the magnetic field and with the detector material, handling multiple scattering, energy loss and much more. This is the most extensive and CPU intensive for the calorimeters, where very many particles are produced in the hadronic/electromagnetic showers. ATLFAST2 has been designed to be able to produce large numbers of models/events with less computing power than would be needed for the full GEANT4 simulation of the whole detector.

The signal and background datasets were simulated with a center of mass energy of $\sqrt{s} =$

10 TeV. During the year of 2009 the energy at which the LHC was going to collide once restarted after the accident of 19 September 2008 was not clear. Hence the choice of the collaboration is somewhat off from the mark of $\sqrt{s} = 7$ TeV at which the LHC started colliding protons in 2010. Chapter 7 will deal with applying the method described in this chapter to the data measured in 2010. The next two sections will describe the signal and background MC samples that were used in this analysis.

6.2.1 Backgrounds

Different MC generators were used for different background samples. This was done in an attempt to optimize the reliability of the estimate and cross check the results for the Standard Model samples.

Top quark production One of the two largest backgrounds for one lepton SUSY searches is $t\bar{t}$ pair production. As the top quark almost always decays into a W and a bottom quark, the signature of a $t\bar{t}$ decay is determined by the decay channels of the two W bosons. In the case of the *fully hadronic* $t\bar{t}$ both Ws decay into a pair of quarks. This case is not interesting from the point of view of one lepton analyses, so we will not mention it again. The other two cases of $t\bar{t}$ decay are of great interest. If one of the Ws decays into a lepton and a neutrino, we speak of *semileptonic* ($l\nu qq$) decays, and if both Ws decay leptonically we speak of *dileptonic* ($l\nu l\nu$) decay.

As $t\bar{t}$ pair production at the LHC can occur well beyond the immediate mass threshold we need to use the next-to-leading order (NLO) generator MC@NLO [53, 54] to accurately describe the hard interaction for $t\bar{t}$ production with initial and final state radiation (ISR/FSR) of hard jets. Parton showering and fragmentation are performed by the HERWIG [55, 56] program. The cross section for the joint semileptonic and dileptonic $t\bar{t}$ sample as calculated by MC@NLO is 217 pb [57] as can be seen in Table 6.1.

For more than two additional hard partons the matrix elements (ME) for NLO $t\bar{t}$ production are too difficult to calculate. Another approach is to use the ALPGEN generator [58] that can calculate ME up to 6 extra partons, but at Leading Order (LO). Here one has to compensate for the overlap between jets created by matrix elements on one side and by parton showering on the other. The technique that accomplishes this is called MLM-matching [59]. The MLM matching procedure starts from generating events using the ME with n extra partons. Each parton has to meet a minimum transverse momentum p_T^{min} requirement, and a minimum separation ΔR^{min} from the other partons. A showering algorithm is applied to each outgoing parton, without a veto on hard emission. Jets are created from the showering output using a cone jet algorithm with radius ΔR^{min} and $p_T > p_T^{min}$. The jets found using this algorithm are matched to the partons from the ME calculation. If m partons are not matched to a jet, this topology belongs to the $n - m$ ME parton sample and it is rejected. Likewise, if m jets are not matched to a parton it belongs to the $n + m$ ME parton sample and is rejected. Thus double counting caused by showering is avoided. To get an inclusive sample including all topologies with extra jets, the last rejection is skipped for the highest multiplicity sample if the extra jets are softer then the ME jets. We mostly use the ALPGEN $t\bar{t}$ events for a study of systematic uncertainties. The cross sections for all the separate ALPGEN samples given in Table 6.1 are calculated after MLM-matching, i.e. the ALPGEN generator cross section is multiplied with MLM-matching efficiency to get the final quoted cross section. The cross

Process	Generator	Cross Section (pb)
$t\bar{t}$ $(l\nu l\nu + l\nu qq)$	MC@NLO	217
$t\bar{t}$ $(l\nu l\nu)$ + 0 partons	ALPGEN	12.7
$t\bar{t}$ $(l\nu l\nu)$ + 1 parton	ALPGEN	13.7
$t\bar{t}$ $(l\nu l\nu)$ + 2 partons	ALPGEN	9.36
$t\bar{t}$ $(l\nu l\nu)$ + 3 partons	ALPGEN	7.06
$t\bar{t}$ $(l\nu qq)$ + 0 partons	ALPGEN	51.8
$t\bar{t}$ $(l\nu qq)$ + 1 parton	ALPGEN	57.1
$t\bar{t}$ $(l\nu qq)$ + 2 partons	ALPGEN	38.3
$t\bar{t}$ $(l\nu qq)$ + 3 partons	ALPGEN	27.6

Table 6.1: Cross sections including NLO k-factor for simulated $t\bar{t}$ processes with different generators. The cross sections for processes generated with ALPGEN are given after MLM matching.

section is compensated for the difference between the NLO and LO calculation by multiplying the LO cross section with a calculated k-factor, scaling it to the NLO value.

W+jets Another important background for SUSY searches with one lepton is production of W-bosons with associated jets. The number of jets in an event is an important selection criterion for this analysis, thus it is very important to correctly produce the kinematics of the additional jets. The ALPGEN generator was chosen to calculate the exact matrix elements for multiparton hard processes, which was interfaced to HERWIG [55, 56] for showering and hadronisation. Alike to the $t\bar{t}$ ALPGEN samples, also here we have to apply the MLM-matching technique, dividing the phase space between matrix elements for hard jets and parton showering only for soft jets, to prevent double counting. Contributions from processes with associated parton multiplicities between zero and five were summed to produce the complete W+jets sample. The cross sections after applying the MLM-matching technique for all lepton flavor separated processes are given in Table 6.2. For the W+jets a NNLO k-factor, calculated using the FEWZ program, was applied to the LO cross section.

W + $\mathbf{b\bar{b}}$+jets The W+jets processes simulated by ALPGEN as described above only take into account light flavor (*u, d, s* and *c* quarks) jets. A separate ALPGEN process takes care of the $W+b\bar{b}$+jets production. Although these processes have relatively low cross sections, they must be considered to correctly estimate the total cross section and if b-tagging is to be used. A small overlap is expected between the $W + b\bar{b}$+jets and W+light jets samples, as the latter may contain $b\bar{b}$ pairs generated by parton showering. This small amount of double counting is minimized by the choice of the generator level cuts [57].

QCD multijet Although the requirement of a single isolated lepton and missing transverse energy strongly suppresses QCD multijet events, the cross sections of these processes are orders of magnitude higher then the processes involving top quarks and W-bosons, and as such still have to be taken into account. ALPGEN was again the generator of choice for it can calculate matrix elements of events with up to 6 partons in the final state. The generation of QCD events was split according to the quark flavor (b-quarks or light quarks) and the transverse momentum of the leading jet, to be able to produce useful amounts of integrated luminosity. The lowest produced p_T of the leading jet was set at 35 GeV, due to practical limitations

Process	Generator	Cross Section (pb)
W(e ν) + 0 partons	ALPGEN	$12.3 \cdot 10^3$
W(e ν) + 1 parton	ALPGEN	$2.6 \cdot 10^3$
W(e ν) + 2 partons	ALPGEN	$8.3 \cdot 10^2$
W(e ν) + 3 partons	ALPGEN	$2.4 \cdot 10^2$
W(e ν) + 4 partons	ALPGEN	67.7
W(e ν) + 5 partons	ALPGEN	19.9
W($\mu\nu$) + 0 partons	ALPGEN	$12.3 \cdot 10^3$
W($\mu\nu$) + 1 parton	ALPGEN	$2.6 \cdot 10^3$
W($\mu\nu$) + 2 partons	ALPGEN	$8.3 \cdot 10^2$
W($\mu\nu$) + 3 partons	ALPGEN	$2.4 \cdot 10^2$
W($\mu\nu$) + 4 partons	ALPGEN	67.7
W($\mu\nu$) + 5 partons	ALPGEN	19.9
W($\tau\nu$) + 0 partons	ALPGEN	$12.3 \cdot 10^3$
W($\tau\nu$) + 1 parton	ALPGEN	$2.6 \cdot 10^3$
W($\tau\nu$) + 2 partons	ALPGEN	$8.3 \cdot 10^2$
W($\tau\nu$) + 3 partons	ALPGEN	$2.4 \cdot 10^2$
W($\tau\nu$) + 4 partons	ALPGEN	67.7
W($\tau\nu$) + 5 partons	ALPGEN	19.9
W($b\bar{b}$) + 0 partons	ALPGEN	6.2
W($b\bar{b}$) + 1 parton	ALPGEN	6.1
W($b\bar{b}$) + 2 partons	ALPGEN	3.5
W($b\bar{b}$) + 3 partons	ALPGEN	2.0

Table 6.2: Cross sections for simulated ALPGEN W+jets processes including NNLO k-factors, which were calculated for an inclusive W sample. The cross sections are given after MLM-matching was applied.

related to cross section of lower p_T samples. The produced QCD samples with corresponding cross sections and allowed leading jet transverse momenta are given in Table 6.3.

6.2.2 Signal generation and mSUGRA grid

In order to cover a large parameter space and different phenomenologies a mSUGRA grid was produced [60]. The grid was made in “radial coordinates”, i.e. points on outgoing radial lines in the $(m_0, m_{1/2})$ plane for $\tan\beta = 10$ and 50. Lines with different slopes in the $(m_0, m_{1/2})$ plane were produced. The other mSUGRA parameters are set to the same values as in [31], i.e. $A_0 = 0\,\mathrm{GeV}, \mu > 0$. HERWIG 6.5 [55, 56] was used to generate SUSY events, based on SUSY spectra generated with ISAJET 7.75 [61].

In our analysis we used approximately 60 models produced in the mSUGRA grid with m_0 values ranging from 50 to 2000 GeV and $m_{1/2}$ values in-between 100 and 450 GeV. The produced SUSY points cross sections extend from $\sim 0.3\,\mathrm{pb}$ for high mass SUSY to approximately 500 pb for the lowest mass SUSY point.

Besides the grid of mSUGRA points, a set of benchmark SUSY points were chosen by ATLAS, taking the principle that the predicted cosmological relic density of neutralinos should be consistent with the observed density of cold dark matter as a guidance. In the mSUGRA scenario this is possible only in restricted regions of the parameter space, hence only 7 points have been chosen [31] that represent different phenomenologies. One of these, known as SU4, is a low mass SUSY point with a next-to-leading order cross section of 402 pb, while the other 6 points have next-to-leading order cross sections of the order 3-30 pb. The benchmark points SU3 and SU4 which will be used for comparison in this chapter have the mSUGRA parameter values:

- SU3: $m_0 = 100\,\mathrm{GeV}$, $m_{1/2} = 300\,\mathrm{GeV}$, $A_0 = 300\,\mathrm{GeV}$, $\tan\beta = 6$, $sgn(\mu) = +$
- SU4: $m_0 = 200\,\mathrm{GeV}$, $m_{1/2} = 160\,\mathrm{GeV}$, $A_0 = -400\,\mathrm{GeV}$, $\tan\beta = 10$, $sgn(\mu) = +$

Both the grid and the benchmark samples were used to define the SUSY ansatz described in section 6.4.4, and to test our method on different realizations of the mSUGRA model.

6.3 Event Selection

Our event selection rests on two principles: we want to keep the highest possible SUSY signal significance in the signal region, while keeping enough of the backgrounds in the control region to correctly estimate their fractions and shapes. The exception to the second principle is the uncertain contribution from QCD multijet events that we try to suppress as much as possible. The event selection criteria can be summarized by:

1. Exactly one isolated muon [1] with $p_T > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$, satisfying identification criteria described in section 3.
2. Second lepton veto, no additional isolated leptons, either electrons or muons, with $p_T > 10\,\mathrm{GeV}$.

---

[1] Due to practical constraints we focus on events with one isolated muon only, though the method described here has been shown to work on events with one isolated electron [31].

Process	Generator	Cross Section (pb)	min	-	max p_T (GeV)
QCD(light jet) + 2 partons	ALPGEN	$3.01 \cdot 10^7$	35	-	70
QCD(light jet) + 3 partons	ALPGEN	$9.84 \cdot 10^6$	35	-	70
QCD(light jet) + 4 partons	ALPGEN	$1.49 \cdot 10^6$	35	-	70
QCD(light jet) + 5 partons	ALPGEN	$2.49 \cdot 10^5$	35	-	70
QCD(light jet) + 2 partons	ALPGEN	$1.12 \cdot 10^6$	70	-	140
QCD(light jet) + 3 partons	ALPGEN	$1.49 \cdot 10^6$	70	-	140
QCD(light jet) + 4 partons	ALPGEN	$5.52 \cdot 10^5$	70	-	140
QCD(light jet) + 5 partons	ALPGEN	$1.90 \cdot 10^5$	70	-	140
QCD(light jet) + 2 partons	ALPGEN	$3.19 \cdot 10^4$	140	-	280
QCD(light jet) + 3 partons	ALPGEN	$6.55 \cdot 10^4$	140	-	280
QCD(light jet) + 4 partons	ALPGEN	$4.90 \cdot 10^4$	140	-	280
QCD(light jet) + 5 partons	ALPGEN	$4.25 \cdot 10^3$	140	-	280
QCD(light jet) + 6 partons	ALPGEN	$1.16 \cdot 10^4$	140	-	280
QCD(light jet) + 2 partons	ALPGEN	$7.50 \cdot 10^2$	280	-	∞
QCD(light jet) + 3 partons	ALPGEN	$1.95 \cdot 10^3$	280	-	∞
QCD(light jet) + 4 partons	ALPGEN	$2.15 \cdot 10^3$	280	-	∞
QCD(light jet) + 5 partons	ALPGEN	$1.39 \cdot 10^3$	280	-	∞
QCD(light jet) + 6 partons	ALPGEN	$9.73 \cdot 10^2$	280	-	∞
QCD($b\bar{b}$) + 0 partons	ALPGEN	$1.38 \cdot 10^5$	35	-	70
QCD($b\bar{b}$) + 1 parton	ALPGEN	$1.94 \cdot 10^5$	35	-	70
QCD($b\bar{b}$) + 2 partons	ALPGEN	$5.38 \cdot 10^4$	35	-	70
QCD($b\bar{b}$) + 3 partons	ALPGEN	$1.35 \cdot 10^4$	35	-	70
QCD($b\bar{b}$) + 0 partons	ALPGEN	$5.40 \cdot 10^3$	70	-	140
QCD($b\bar{b}$) + 1 partons	ALPGEN	$2.72 \cdot 10^4$	70	-	140
QCD($b\bar{b}$) + 2 partons	ALPGEN	$1.86 \cdot 10^4$	70	-	140
QCD($b\bar{b}$) + 3 partons	ALPGEN	$9.46 \cdot 10^3$	70	-	140
QCD($b\bar{b}$) + 0 partons	ALPGEN	$1.48 \cdot 10^2$	140	-	280
QCD($b\bar{b}$) + 1 parton	ALPGEN	$1.08 \cdot 10^3$	140	-	280
QCD($b\bar{b}$) + 2 partons	ALPGEN	$1.43 \cdot 10^3$	140	-	280
QCD($b\bar{b}$) + 3 partons	ALPGEN	$1.02 \cdot 10^3$	140	-	280
QCD($b\bar{b}$) + 4 partons	ALPGEN	$7.07 \cdot 10^2$	140	-	280
QCD($b\bar{b}$) + 0 partons	ALPGEN	3.24	280	-	∞
QCD($b\bar{b}$) + 1 parton	ALPGEN	25.2	280	-	∞
QCD($b\bar{b}$) + 2 partons	ALPGEN	50.0	280	-	∞
QCD($b\bar{b}$) + 3 partons	ALPGEN	52.9	280	-	∞
QCD($b\bar{b}$) + 4 partons	ALPGEN	55.5	280	-	∞

Table 6.3: Leading Order cross sections for simulated ALPGEN QCD multijet processes, that were split according to quark flavor and p_T of the leading jet. The last two columns give the minimal and maximal transverse momentum of the leading jet as was used for the separation of the samples. The cross sections are given after MLM-matching was applied.

3. $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2 \times M_{\mathrm{eff}}$.
4. At least four jets with $|\eta| < 2.5$ and $p_T > 20\,\mathrm{GeV}$, out of which the leading jet has to have $p_T > 80\,\mathrm{GeV}$ and the two sub-leading jets must have $p_T > 40\,\mathrm{GeV}$.

The first cut defines the one-lepton analysis, while the second ensures that the overlap with the two-lepton analyses is absent, as well as cutting the dileptonic $t\bar{t}$ background. Cuts 3 and 4 both strongly reduce the Standard Model backgrounds, while keeping much of the SUSY signal intact. Our choice for the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold of 40 GeV is somewhat lower than the baseline ATLAS SUSY cut, as we are focused not only on maximizing the signal significance but also the event count of the backgrounds. For the same reason we do not use a $M_{\mathrm{T}} > 100\,\mathrm{GeV}$ cut usually used by other analyses studying SUSY in the one-lepton case.

A study has been performed[31] to conclude that the $E_{\mathrm{T}}^{\mathrm{miss}} > 0.2 \times M_{\mathrm{eff}}$ cut is effective in getting rid of most SM backgrounds, specifically the multijet QCD events. For SUSY events the values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{eff} are quite correlated, while much less so for backgrounds. Two other variables can be used to suppress the QCD background, transverse sphericity (S_T), and azimuthal angle difference between jets and missing transverse energy ($\Delta\phi(jet, E_{\mathrm{T}}^{\mathrm{miss}})$). The first assumes that the heavy SUSY particles produced in the initial interaction are almost at rest, hence their decay products would be distributed isotropically, while for the QCD events the direction of the two partons from the hard scattering provides a privileged direction. The second is based on the fact that for QCD events $E_{\mathrm{T}}^{\mathrm{miss}}$ is closely associated to one of the leading jets, either fake from mismeasurement or real from decay of shower particles, thus requiring $\Delta\phi(jet, E_{\mathrm{T}}^{\mathrm{miss}}) > 0.2$ for the leading three jets would suppress the QCD background. However these two variables are not used in our analysis, as we already get enough QCD suppression from the event selection criteria described above.

Table 6.4 summarizes the event counts before and after event selection for an integrated luminosity of 200 pb^{-1}. The W+jets background is the sum of all the ALPGEN produced samples described earlier. The number of events before selection counts only events in a QCD sample is obtained by combining the samples of Table 6.3 for which the leading jet has a $p_T > 280\,\mathrm{GeV}$. The other QCD samples did not contribute to the number of events after selection at all. As can be seen in the table, after our event selection the QCD background can be considered negligible in comparison with the other background samples. It is possible that the MC will not describe actual ATLAS data well; in that case a dedicated data-driven QCD background study can be done that is beyond the scope of this thesis.

Sample	N_{events} Before	N_{events} After
$t\bar{t}$ $(l\nu qq)$	32328	471
$t\bar{t}$ $(l\nu l\nu)$	8768	102
W+jets	7946800	599
QCD	1481000	7
SU4	21480	975
SU3	1092	27

Table 6.4: Event counts before and after selection for an integrated luminosity of 200 pb^{-1} of the backgrounds and two SUSY models, one high mass (SU3) and one low mass (SU4).

The MC@NLO $t\bar{t}$ sample here is split into its dileptonic and semileptonic constituents, on the basis of information of the simulated matrix elements that can be accessed. The reasons for this separation will become evident to the reader in the discussion of the analysis method (section 6.4). Surprisingly enough the dileptonic $t\bar{t}$ remains a sizable component of the SM background, even after the 2nd lepton veto. One of the leptons from the W decay is in this case not identified as an electron or muon. A study showed that these events can be classified into three categories: half contain a $W \rightarrow \tau\nu$ decay where the tau decays hadronically, in 20% of these events the lepton is misidentified as a jet, for the remaining 30% of the events the lepton is either lost inside a jet or falls outside the p_T or η acceptance.

For comparison we include two ATLAS SUSY benchmark points in Table 6.4, a high mass SUSY point SU3 and a low mass SUSY point SU4. For an integrated luminosity of 200 pb^{-1}, it would be very demanding to find high mass SUSY such as SU3 in LHC data with only these cuts. On the other hand the SU4 model still retains many events after our selection. Even though the event selection efficiency is not very high, it is higher than for the SM background: the ratio SU4/$t\bar{t}$ for instance goes from 0.66 to 2.1. Later in this chapter we will define a signal region that further improves this ratio.

Trigger efficiency

Studies were performed to assess the trigger efficiencies for the initial LHC running scenario [60, 31]. For the channels implemented in the simulation of the 10 TeV center-of-mass energy samples the combined jet ($p_T > 70\,\mathrm{GeV}$)+ ($E_{\mathrm{T}}^{\mathrm{miss}} > 70\,\mathrm{GeV}$) j70xE70 trigger yields an efficiency of $> 99\%$ for SUSY events. A trigger on isolated muons with p_T of 6 GeV, (`mu6`), also was found to be over 80% efficient past the turn-on curve. In our event selection we require muons with $p_T > 20\,\mathrm{GeV}$, so we are safely in the `mu6` trigger plateau. For simplicity in the rest of the study we assume a trigger efficiency of 100%.

6.4 Fit method

This section will describe a model for each of the SM backgrounds, based on the simulated samples described in the previous section. The first model will ignore correlations between observables. It will however become clear that these correlations are significant, and the model will be adapted to incorporate them.

The models of semileptonic $t\bar{t}$ and W+jets have been adapted. The semileptonic $t\bar{t}$ model fits only those events for which the hadronic top has been correctly reconstructed (top peak events). All other semileptonic top events will be added to the W+jets sample in a combined W+top combinatorics sample. How this splitting was performed will be described in detail in this chapter.

Finally, a Ansatz model which takes into account the contamination of mSUGRA events in our control region will be developed, and compared to a wide range of points on the mSUGRA grid.

6.4.1 The uncorrelated background model

The backgrounds to SUSY in the one lepton channel are those processes that mimic the SUSY signature of high missing transverse energy combined with a lepton and energetic jets, such as $t\bar{t}$ and W+jets events. It is not possible to distinguish between these processes and SUSY production on an event by event basis. Therefore we construct a model that combines specific physics features of these backgrounds, such as mass peaks or phase space cutoffs.

Name	Description	Comment
TT	Leptonic $t\bar{t}$	generated using MC@NLO/ALPGEN
$T1$	Semileptonic $t\bar{t}$	Split from TT on generator level
$T2$	Dileptonic $t\bar{t}$	Split from TT on generator level
TP	Top peak	Split from $T1$ using method of section 6.4.3.
TC	Top combinatorics	Split from $T1$ using method of section 6.4.3.
W	W+jets	Generated using ALPGEN
CW	W plus TC	

Table 6.5: The different definitions of the background components.

We model the backgrounds by their behavior in three observables, $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj}, which are defined in chapter 3. These observables combine features specific to SUSY, hard $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} spectra, with features that are unique to the SM backgrounds. These characteristics include the top mass peak in M_{jjj} for semileptonic $t\bar{t}$, and a W-mass cutoff in M_{T} for W+jets and semileptonic $t\bar{t}$ events. We have not been able to define an observable that shows a unique signature for dileptonic $t\bar{t}$. Our model depends on a combination of the behavior of dileptonic $t\bar{t}$ in all three observables.

These are the physical backgrounds, but for the sake of fit stability it is advantageous to make one change to our background source definition. When looking at the M_{jjj} distribution of the semileptonic top sample, a pronounced Gaussian peak can be seen where the three jets with the highest $\sum \vec{p}_T$ form the mass of the top. Besides the peak there is a large exponentially decaying component where the three jets that were selected do not correctly identify the original top quark. This we call the *top combinatorics* background component. Since a top mostly decays to a W, these events where the top was not recognized will look very similar to W+jets events. It turned out to be very hard to separate these two types of events. However, we are not primarily interested in determining the relative yield of $t\bar{t}$ and W. Therefore we split the top sample into a 'top peak' and a 'top combinatorics' sample, and add the latter to the W+jets sample. Table 6.5 summarized the definitions of the different background samples used in this analysis. We will now try to model the top peak (TP), W+combinatorics (CW), and the dileptonic $t\bar{t}$ sample ($T2$).

In order to model the TP and CW contribution, we want to have MC simulated samples of these two backgrounds. That means that we have to find a way to separate the semileptonic $t\bar{t}$ sample in two parts. This separation of top peak and top combinatorics events must happen in an unbiased way, which does not introduce method dependencies in our final result, which is discussed in more detail in section 6.4.3. The separated sample serves two purposes: it can be used to get the generic shape of the TP and CW background distributions, and to get an estimate of the total number of TP and CW events that are expected after the selection cuts that were defined previously.

Our model is based on one-dimensional (1D) probability density functions (PDF's), one PDF for each background and each observable. In the absence of correlations we construct a 3D PDF for each background like this:

$$F_{BG}^{3D}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = F_{BG}^{1D}(E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{BG}^{1D}(M_{\mathrm{T}}) \times F_{BG}^{1D}(M_{\mathrm{jjj}}) \tag{6.1}$$

To construct the complete model for the all the SM backgrounds we simply add the 3D PDF's together. The total combined model is thus created like this:

$$PDF^{3D}_{comb} = N_{obs}\left\{(1-f_{TT})PDF^{3D}_{CW} + f_{TT}\left[(1-f_{T2})PDF^{3D}_{TP} + f_{T2}PDF^{3D}_{T2}\right]\right\} \tag{6.2}$$

from the W+jets, semileptonic and dileptonic $t\bar{t}$ 3D PDF's. Here f_{T2} denotes the fraction of $T2$ events relative to the number of TP+$T2$ events, and f_{TT} denotes the fraction of TP+$T2$ events relative to the total number of background events. These fractions are parameters of the model. By defining them recursively, as was done in here, we make sure that the combined PDF is always unit normalized, as long has the values of the fractions stay between 0 and 1. This enhances the stability of the fitting procedure.

Before we discuss our model we define two functions which are used widely. The first is called *TTComb*, since it was first used to describe the combinatorics background of the $t\bar{t}$ sample. It is defined as:

$$\text{TTComb}(x;\mu,\sigma,\alpha) = \frac{1}{N}(1 + erf(x;\mu,\sigma)) \times e^{-\alpha x}, \tag{6.3}$$

with N the normalization constant, x the observable and μ, σ and α the parameters of this PDF. The shape of this PDF can be seen in Figure 6.3, where also the effect of changing the parameters is shown. The TTComb is an exponential at high values of x, where at low values the $erf(x)$ function gives a smooth cutoff to zero. Width and position of the cutoff are controlled by the σ and μ parameters, while the exponential decay is determined by the α parameter.

The second new function that is defined in this analysis is the *MTFunc*. It is used to describe the $M_{\rm T}$ distribution of all SM backgrounds, and it is created to fit the three main features of these distributions: a Gaussian core, a large exponential tail at high $M_{\rm T}$, and a plateau at low $M_{\rm T}$. This function is a sum of two Gaussians and an exponential. Defining the double Gaussian as:

$$\begin{gathered}\text{DG}(x;\mu_{core},\sigma_{core},\mu_{wide},\sigma_{wide},f_{core}) = \\ f_{core}G(x;\mu_{core},\sigma_{core}) + (1-f_{core})G(x;\mu_{wide},\sigma_{wide}),\end{gathered} \tag{6.4}$$

with G denoting a normalized Gaussian PDF, and the fractions such that the normalization of the function can be maintained to unity. The MTFunc is then defined as:

$$\begin{gathered}\text{MTFunc}(x;\mu_{core},\sigma_{core},\mu_{wide},\sigma_{wide},\alpha,f_{peak},f_{core}) = \\ \tfrac{1}{N_{tail}}(1-f_{peak})e^{-\alpha x} + f_{peak}\text{DG}(x;\mu_{core},\sigma_{core},\mu_{wide},\sigma_{wide},f_{core}).\end{gathered} \tag{6.5}$$

Here the exponential has to be normalized explicitly in order for MTFunc to be normalized to 1. The shape of this distribution, and its components can be seen in Figure 6.4.

	$E_{\rm T}^{\rm miss}$	$M_{\rm T}$	$M_{\rm jjj}$
TP	TTComb	MTFunc	Gaussian
CW	f_1TTComb1+$(1-f_1)$TTComb2	MTFunc	TTComb
$T2$	TTComb	MTFunc	TTComb

Table 6.6: The components of the background model. Each 3D background model component consists of three 1D components, one for each observable. The three 3D components are combined as in equation 6.2.

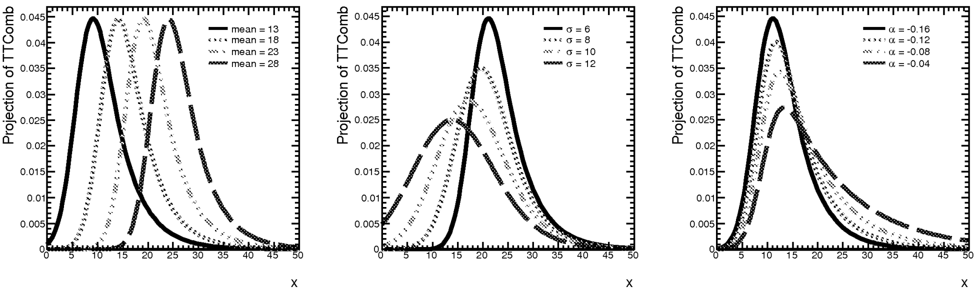


Figure 6.3: Figures showing the shape of the TTComb function for different parameter values. The values of μ, σ and α increase from solid to dashed lines.

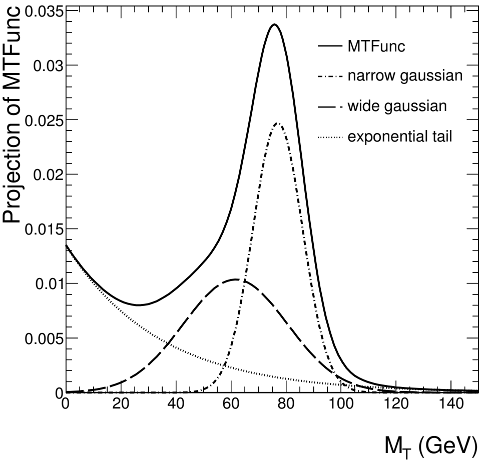


Figure 6.4: The MTFunc distribution and its components.

Table 6.6 shows the models used to describe the three backgrounds. A TTComb is used for the TP and $T2$ $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions. For the CW $E_{\mathrm{T}}^{\mathrm{miss}}$ model a sum of two TTCombs is used to account for a kink in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. There is a subtlety in handling this model in a fit that has to be addressed here: having two TTCombs in one model will confuse the fitting algorithm, if it is unclear which model has the larger μ or σ and which has the smaller. By replacing e.g. μ_{large} with $\mathrm{r}_{\mu} \times \mu_{\mathrm{small}}$, and forcing the ratio r_{μ} to be larger then one, we avoid this ambiguity. The same trick is used for the two Gaussians in MTFunc, where both the μ and the σ of the wide Gaussian are replaced by a ratio term multiplied with respective core Gaussian terms. The M_{jjj} distributions for $T2$ and CW are also described by a TTComb. Finally, the M_{jjj} distribution of the TP sample shows the characteristic Gaussian top mass peak. The definition of the TP model also contains a small TTComb component, to account for the fact that the splitting of TP and TC is not perfect. However, we assume that in the fit of the combined background model to MC data, the small amount of fake TP gets absorbed in the CW fraction, and the TTComb fraction in the TP M_{jjj} model is fixed to 0.

In total this gives 13 parameters for the $E_{\mathrm{T}}^{\mathrm{miss}}$ model, 21 parameters for the M_{T} model, and 8 parameters for the M_{jjj} model. This means the background model has 42 parameters, plus 2 recursive fractions when the three models are combined. A way to reduce the number of parameters is to see which parameters have the same value for different backgrounds, or even within one background model. Which parameters have been merged to form *global* parameters is shown in table 6.7. This reduces the number of shape parameters to 38.

Many of these parameters can be understood to be equal for different backgrounds. As the M_{T} distribution is dominated by the W mass for both TP and CW, they share a few shape parameters. The same holds for the M_{jjj} distribution of CW and $T2$, as both are the product of a random combination of jets. The similarities in the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb widths are most likely artifacts of the cut on the $E_{\mathrm{T}}^{\mathrm{miss}}/M_{\mathrm{eff}}$ ratio.

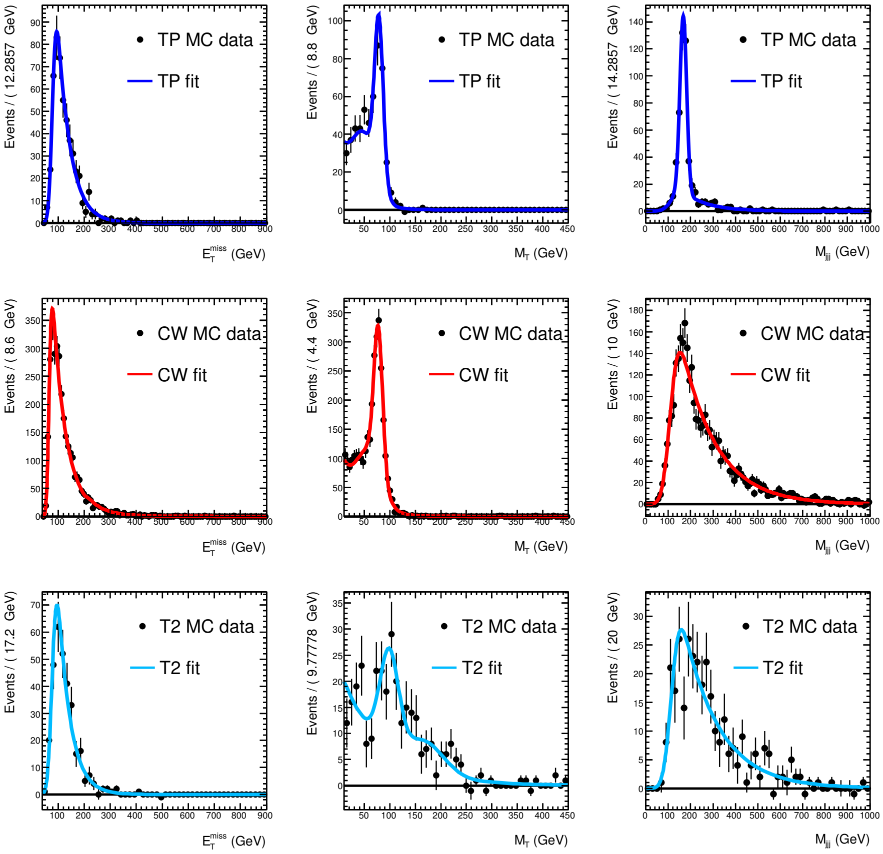


Figure 6.5: 1D projections of all the background models in the three observables. The left column shows the $E_{\mathrm{T}}^{\mathrm{miss}}$, the middle the M_{T} and the right column the M_{jjj} distributions.

The model without correlations is shown in Figure 6.5. The dots represent the simulated events for that specific background which pass all selection criteria described in the previous

Description	Parameter 1	Parameter 2
Width of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb	`TP_etmiss_sigma`	`T2_etmiss_sigma`
Relative size of the core Gaussian	`TP_mtrans_fcore`	`CW_mtrans_fcore`
Width of the narrow M_{T} Gaussian	`TP_mtrans_sigma_core`	`CW_mtrans_sigma_core`
Width of the M_{jjj} TTComb	`CW_mjjj_sigma`	`T2_mjjj_sigma`
Width of both *CW* $E_{\mathrm{T}}^{\mathrm{miss}}$ TTCombs	`CW_etmiss_sigma_1`	`CW_etmiss_sigma_2`

Table 6.7: Sets of parameters that are combined into one global parameter. The naming convention of the parameters is as follows: `<model>_<observable>_<parameter>`, thus `CW_etmiss_sigma` means the width of the *CW* $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

section. The solid line represent the 1D projection of the 3D background PDF. Each of the three background PDFs was fitted to simulated data for that specific background. These plots show a nice resemblance of our model with the simulated data. But since they are projections, they may hide correlations between observables. In the next section, it will become clear that due to such correlations the shape of the M_{T} and M_{jjj} distribution can change depending on the value of $E_{\mathrm{T}}^{\mathrm{miss}}$.

6.4.2 The correlated background model

So far we have been considering factorizable models or models that can be described by a simple product of one-dimensional models. The observables have been chosen to be as uncorrelated as possible, but the model parameters in observable M_{T} and M_{jjj} are correlated to observable $E_{\mathrm{T}}^{\mathrm{miss}}$. These correlations can be described by conditional PDF's as discussed in section 5. We will now discuss building of one of these conditional PDF's for our model, which will be followed by a short discussion of all the conditional parameters introduced in the *correlated model*.

We focus here on the shape of *CW* in $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{jjj}. If no correlation is present in the PDF $CW(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{jjj}})$ then the shape of $CW(M_{\mathrm{jjj}})$ is the same for each value of $E_{\mathrm{T}}^{\mathrm{miss}}$. We can check if this is correct by looking at the *CW* M_{jjj} distribution in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$, as shown in Figure 6.6. If no correlation between M_{jjj} and $E_{\mathrm{T}}^{\mathrm{miss}}$ exists, the shape parameters of the M_{jjj} model should be approximately constant as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$.

The shape of *CW* M_{jjj} distribution is described by the TTComb function. The TTComb function has three parameters α, μ, σ. Out of these we pick α as an example, which is the parameter of the exponentially decaying tail of the TTComb. If we slice the *CW* M_{jjj} distribution in bins of missing transverse energy, as shown in Figure 6.6, we can fit our TTComb model to each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin separately as shown by the green dashed line in each subfigure. If we now take the value and error of parameter α for each separate bin, we can see if and how it evolves with increasing $E_{\mathrm{T}}^{\mathrm{miss}}$. This evolution is shown in Figure 6.7(a) by the black markers. Clearly, the shape of the M_{jjj} model changes with changing $E_{\mathrm{T}}^{\mathrm{miss}}$.

We can now parametrize the evolution of $\alpha(E_{\mathrm{T}}^{\mathrm{miss}})$ with an analytical function. The evolution can be described using the error function (*erf*) as follows:

$$\alpha(E_{\mathrm{T}}^{\mathrm{miss}}) = A \times \mathrm{erf}((E_{\mathrm{T}}^{\mathrm{miss}} - \mu)/\sigma) + B, \tag{6.6}$$

where one PDF parameter α is replaced by four conditional parameters A, B, μ and σ. Now we

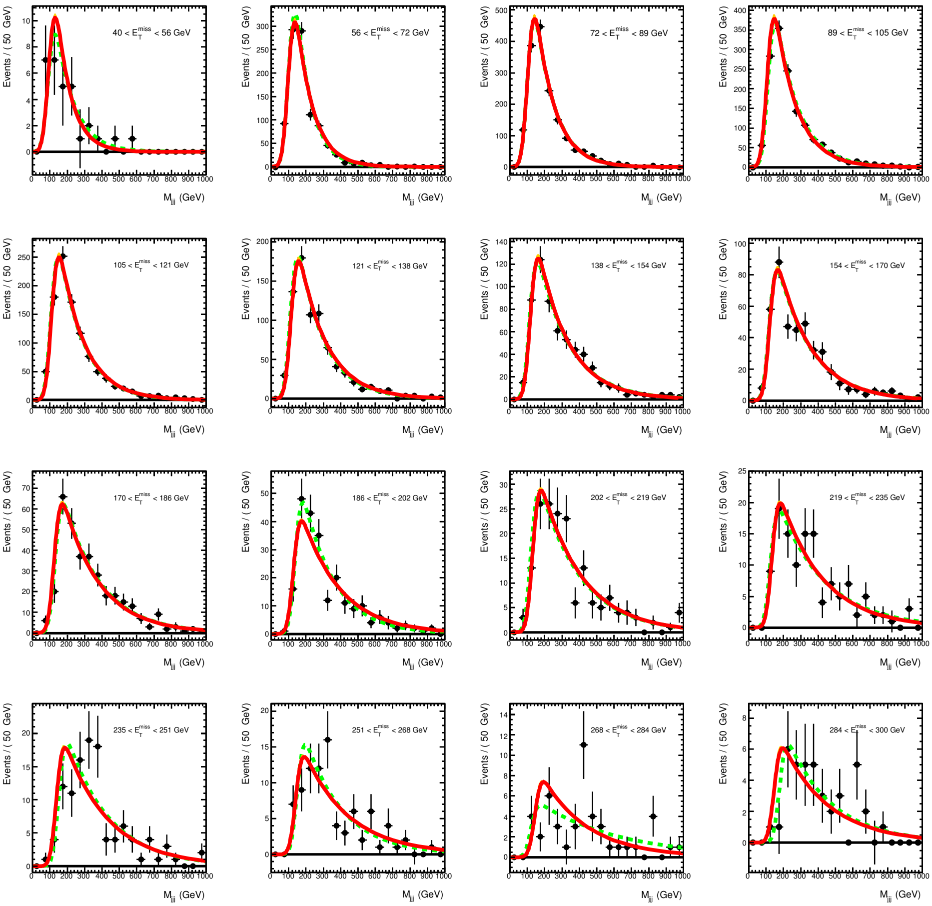
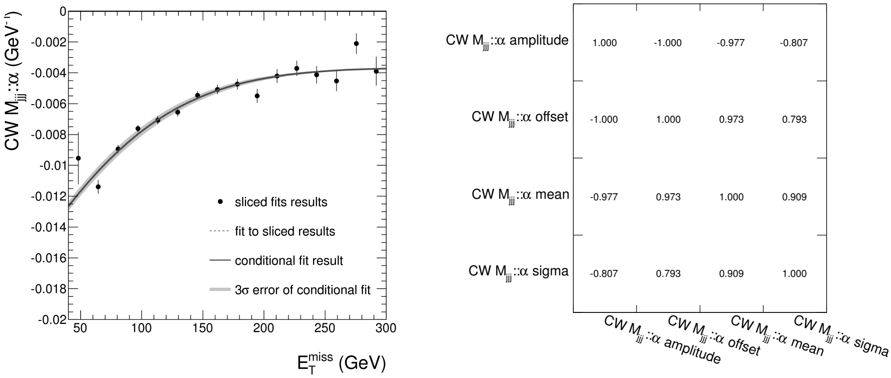


Figure 6.6: $M_{\rm jjj}$ distribution of the CW sample sliced in $E_{\rm T}^{\rm miss}$. A separate bin fit is given by the dashed line, while the fit result for the correlated model fit is given by the solid line. The propagated errors of the correlated model fit are given in light grey with 3 standard deviations.

can rewrite PDF $CW(E_{\rm T}^{\rm miss}, M_{\rm jjj})$ as $CW_1(M_{\rm jjj}|E_{\rm T}^{\rm miss}) \times CW_2(E_{\rm T}^{\rm miss})$ by substituting parameter α with $\alpha(E_{\rm T}^{\rm miss})$ in $CW_1(M_{\rm jjj}|E_{\rm T}^{\rm miss})$. However if we fit this function to the evolution of α, as shown by the green dashed line in Figure 6.7(a), we see that some of these parameters are highly correlated. The correlation matrix of this fit to sliced results can be seen in Figure 6.7(b), from which we learn that many of these parameters are redundant. These strong correlations can be dissolved by setting some of these parameters constant. We have chosen to set the parameters A, B and μ constant. This leaves us with a function $\alpha(E_{\rm T}^{\rm miss})$ that is dependent only on one parameter, σ. Hence we replace the parameter α in our model by the $E_{\rm T}^{\rm miss}$ dependent function $\alpha(E_{\rm T}^{\rm miss})$ without increasing the total number of parameters.

If we now redo the full CW fit in three dimensions without slicing it, we can study whether our function describes the evolution correctly. This can be seen by comparing the fits to data



(a) (b)

Figure 6.7: Correlation (a) between the $CW\ M_{\mathrm{jjj}} :: \alpha$ parameter and $E_{\mathrm{T}}^{\mathrm{miss}}$. In black dots are results of separate bin fits, which were fitted by a dashed line. The solid line is the result of the correlated model to the full CW sample. Correlation matrix (b) of the fit to separate bin results shown by the line in the left plot.

in separate $E_{\mathrm{T}}^{\mathrm{miss}}$ bins, shown by the green dashed line in figures 6.6 and 6.7(a), to the solid line that shows the projected result of the correlated model fit in each of the $E_{\mathrm{T}}^{\mathrm{miss}}$ bins. What we conclude is that the evolution of α is correctly described by the correlated model, while the parameter that is left floating, *sigma*, will be able to account for a possible deviation between data and simulation.

We have only found correlations between parameters in $M_{\mathrm{T}}/M_{\mathrm{jjj}}$ models and the observable $E_{\mathrm{T}}^{\mathrm{miss}}$. No correlations were found between parameters of M_{T} models and M_{jjj} or vice versa. It turns out that for the TP model, the M_{T} shape depends on $E_{\mathrm{T}}^{\mathrm{miss}}$, and for the CW and $T2$ models, both the M_{T} and the M_{jjj} shapes depend on $E_{\mathrm{T}}^{\mathrm{miss}}$. Hence, the general shape of the 3D PDF for the TP background will be :

$$F_{TP}^{3D}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = F_{TP}^{1D}(E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{TP}^{1D}(M_{\mathrm{T}}|E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{TP}^{1D}(M_{\mathrm{jjj}}) \tag{6.7}$$

and for CW and $T2$:

$$F_{CW/T2}^{3D}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = F_{CW/T2}^{1D}(E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{CW/T2}^{1D}(M_{\mathrm{T}}|E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{CW/T2}^{1D}(M_{\mathrm{jjj}}|E_{\mathrm{T}}^{\mathrm{miss}}) \tag{6.8}$$

with $F(x|y)$ denotes a PDF in x which is conditional in y, as discussed in section 5.

In table 6.8 we list all the parameters that are replaced as well as the corresponding conditional functions, separated by the background sample. As we explained in our example above, concerning the $CW\ M_{\mathrm{jjj}} :: \alpha$ parameter, some of the conditional parameters are redundant due to high correlations, hence the superscript c in table 6.8 denotes the parameters that were made constant. As before *erf* denotes the error function, while *erfc* is the complementary error function. Although the total number of parameters increases when we replace a parameter by a function of multiple parameters, due to the redundancies the total number of floating parameters in the combined fit does not increase too much.

Finally, an interesting observation that we made during the correlations study is the strong evolution of the M_{T} distribution of dileptonic $t\bar{t}$ background compared to the other samples,

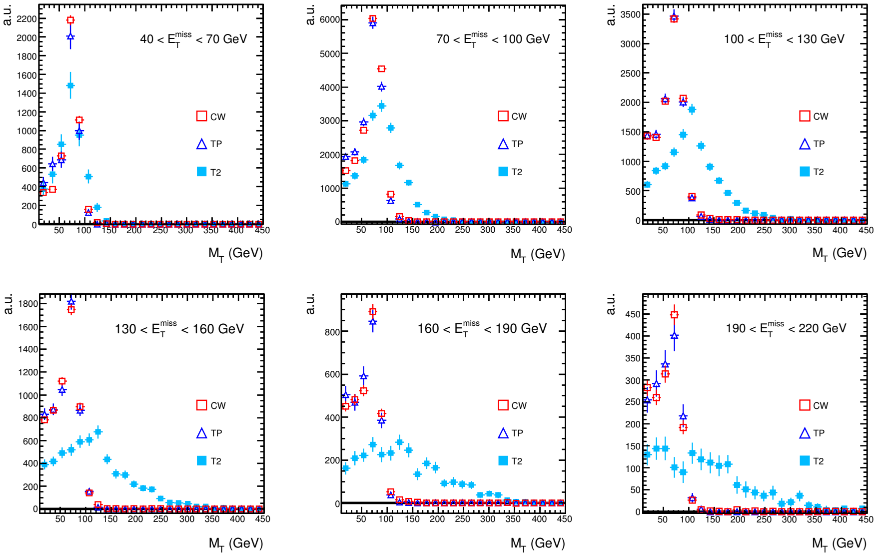


Figure 6.8: Normalized M_{T} distributions of the three main background samples sliced in $E_{\mathrm{T}}^{\mathrm{miss}}$. Dileptonic $t\bar{t}$ distribution has a remarkably pronounced evolution comparing to the other two samples, that serves to distinguish it in a combined fit.

as shown in Figure 6.8. This elusive background has been a source of much trouble for many data-driven methods developed by the ATLAS SUSY working group, but the unique behavior shown in the figure should give our method a handle on its shape and yield.

As mentioned in section 6.4.1 it is advantageous to search for parameters that have similar values, reducing the total number of parameters. We repeat this study after taking correlations into account in our models, to find three conditional parameters in M_{T} that can be shared between the CW and TP models.

6.4.3 Top peak, combinatorics separation

As we stated before we observe two distinct parts in the semileptonic top sample, based on the shape of the M_{jjj} distribution. There is a clear Gaussian peak where the three jets with highest $\sum \vec{p}_T$ have a combined mass around the top mass. Besides the peak there is a large combinatorics background, where the top mass was not reconstructed. Loosely speaking, one might say that for the events that make up the Gaussian peak the correct three jets formed by partons from the hadronically decaying top are selected by the M_{jjj} algorithm. If we want to treat these two parts of the semileptonic top sample separately for the sake of fit stability, we have to make this loose statement more precise.

To get a model of the TP and CW background, a simulated background sample is needed for each background. These two samples are obtained from the semileptonic $t\bar{t}$ sample, which is split in two parts, and the W+jets sample, which is enlarged with the top combinatorics events. The splitting of the semileptonic $t\bar{t}$ sample in an unbiased way is described in this section.

Parameter	Replacing Function
	CW
$M_{\mathrm{T}}::f_{peak}$	$min(1, erf((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^c)/\sigma^c) + b)$
$M_{\mathrm{T}}::\mu$	$b + E_{\mathrm{T}}^{\mathrm{miss}} \times a^c$
$M_{\mathrm{T}}::r_\mu$	$b^c + E_{\mathrm{T}}^{\mathrm{miss}} \times a$
$M_{\mathrm{jjj}}::\alpha$	$A^c \times erf((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^c)/\sigma) + b^c$
$M_{\mathrm{jjj}}::\mu$	$b + E_{\mathrm{T}}^{\mathrm{miss}} \times a$
	TP
$M_{\mathrm{T}}::f_{peak}$	$min(1, erf((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^c)/\sigma) + b^c)$
$M_{\mathrm{T}}::r_\mu$	$max(0.5, b^c + E_{\mathrm{T}}^{\mathrm{miss}} \times a)$
	T2
$M_{\mathrm{T}}::\mu$	$min(250, b + E_{\mathrm{T}}^{\mathrm{miss}} \times a^c)$
$M_{\mathrm{T}}::\sigma$	$min(35, b + E_{\mathrm{T}}^{\mathrm{miss}} \times a^c)$
$M_{\mathrm{T}}::f_{peak}$	$min(1, A^c \times erfc((E_{\mathrm{T}}^{\mathrm{miss}} - \mu)/\sigma^c)$
$M_{\mathrm{T}}::\alpha$	$min(-1 \cdot 10^{-4}, b : +erf((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^c)/\sigma^c) + a^c \times E_{\mathrm{T}}^{\mathrm{miss}}$
$M_{\mathrm{jjj}}::\mu$	$b^c + E_{\mathrm{T}}^{\mathrm{miss}} \times a^c$
$M_{\mathrm{jjj}}::\alpha$	$A^c \times erf((E_{\mathrm{T}}^{\mathrm{miss}} - \mu^c)/\sigma) + b^c$

Table 6.8: Model parameters with corresponding functions that replaced them. Conditional parameters with a superscript c are kept constant in the fit, as they are found to be redundant. In the formulas a denotes a slope, b an offset, A an amplitude, μ and σ are the μ and std. deviation of the Gaussians from which the error functions are derived. Although for the sake of simplicity these variables have the same name in each formula above, they are different parameters of the model.

The splitting of the semileptonic top sample can be done by matching the reconstructed jets that have the highest $\sum \vec{p}_T$ to the three quarks that are the decay products of the hadronically decaying top. Since we are dealing with MC simulation, the momentum information of these quarks is available. If the jets match the quarks, the event is tagged as top peak (TP), if not the event is tagged top combinatorics (TC). However, there are complications.

Since the top is a color carrying particle, top quark production and decay is accompanied by production of associated radiation. The radiation of colored partons or gluons can result in additional hard partons, originating from before the hard interaction (Initial State Radiation, ISR) or originating from after the hard interaction (Final State Radiation, FSR). Both ISR and FSR contribute to the top pair production matrix element. Radiation of a hard parton results in an observable extra jet as this parton hadronizes. More difficult to control however is the soft gluon radiation, where M_{jjj} might be off from m_{top} for such topologies even when the reconstructed tops and the three quarks from top decay are matched.

Ideally one would match all reconstructed jets to the partons that are the decay products of the top, and than determine if these decay products were correctly identified by the M_{jjj} algorithm, whereas for the combinatorics background this is not the case.. However, comparing a reconstructed object such as a jet with a simulated parton is troublesome as all partons radiate and a quarkline or the sum of three quarklines is not an observable as such. To sidestep this in comparing reconstructed jets to simulated particles we use so-called *truth-jets* instead of partons. These objects are made by running the jet algorithm on all observable

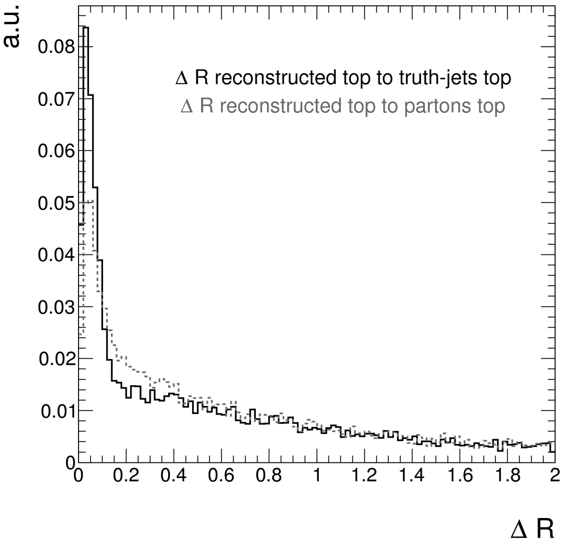


Figure 6.9: Normalized distributions of ΔR between the reconstructed top and the truth jets top (solid line) and between the reconstructed top and the partonic top (dashed line).

Monte Carlo simulation particles after showering and hadronization. This way truth-jets take soft gluon radiation into account and we are no longer strongly dependent on the amount of soft radiation for each event. Truth-jets are what we expect to measure in the limit that we have an ideal detector and assuming that MC fragmentation and hadronization are able to correctly describe nature. [2]

An illustrative example of the difference between truth-jets and partons is shown in Figure 6.9. This figure considers three different ways of reconstructing the top quark in a MC simulated sample. The first, the so-called truth-top is obtained from the vector sum of the 4-vectors of the three top decay products which are stored in the MC-truth information containing the output of the event generator. This output is at parton level, so the three decay products are a b-quark and the two quarks (since we are looking at the hadronically decaying top) from W decay.

Second, the truth-jet top is defined as the vector sum of the three truth-jets as defined above that were best matched in $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ to the three quarks from top decay.

Finally, after detector simulation and reconstruction, the three reconstructed jets with the highest $\sum \vec{p}_T$ form the reconstructed top.

Figure 6.9 shows the normalized ΔR distribution between the reconstructed hadronic top and the truth-jets top as a solid line. Superimposed as a dashed line is the normalized ΔR distribution of the reconstructed hadronic top to the partonic top. What is clear is that for truth-jets there is a more apparent cut-off, while the truth top distribution drops off more fluently. Both distributions have a big tail due to incorrectly reconstructed or mismatched

---

[2] To make sure that we do not have double counting, we do overlap removal between truth-jets and truth-electrons just the way we do it for reconstruction objects.

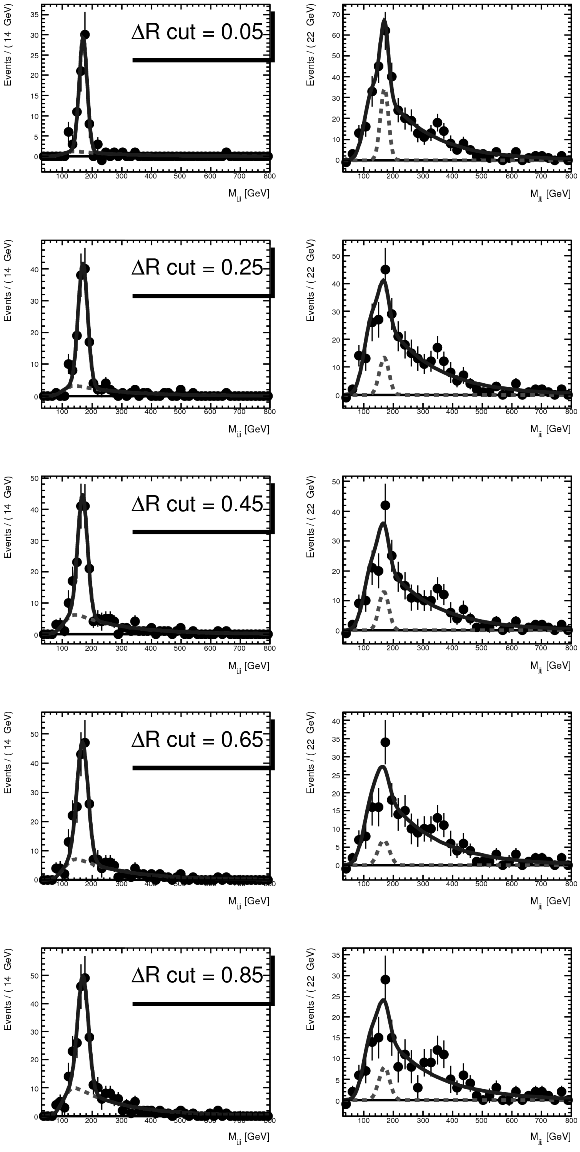


Figure 6.10: The column on the left shows the M_{jjj} distribution of top events tagged as top peak, for a given ΔR cut. The column on the right shows the events tagged as top combinatorics. Each row shows with a solid line TP model to the peak events, and a TC model on the combinatorics events, which were simultaneously fitted to their respective sample. The dashed line shows the fitted fake fraction in each sample.

top quarks, but the soft gluon radiation that perturbs the partons momentum and direction is much less pronounced for truth-jets.

We will use the ΔR between the truth-jets top and reconstructed top to split the semilep-

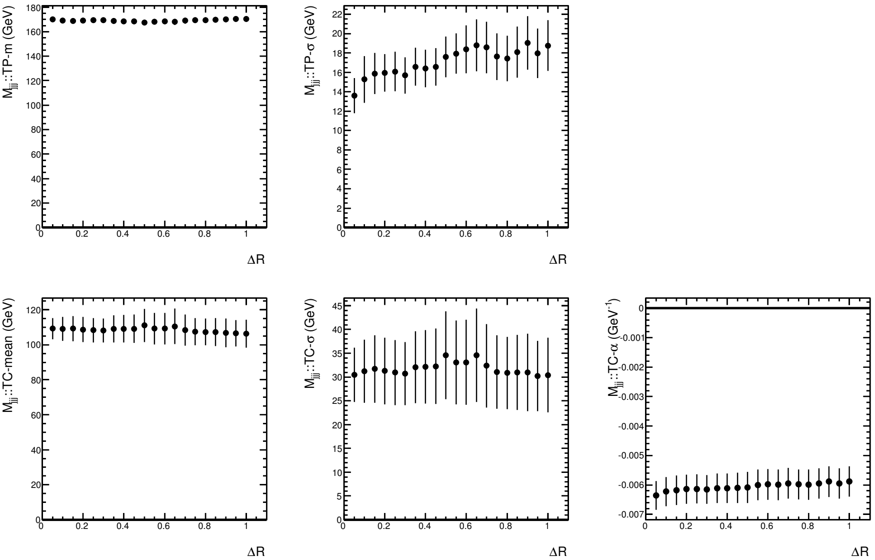


Figure 6.11: The evolution of simultaneous fit parameters as function of ΔR. On top are two parameters describing our top peak model $TP-\mu$ and $TP-\sigma$, and on the bottom are the three parameters of the combinatorics model $TC-\mu$, $TC-\sigma$ and $TC-\alpha$. Within the error the values of all parameters stay constant, giving us confidence that our model of the MC data is correct and unbiased. The small increase in the value of $TP-\sigma$ is understood. The TP distribution gets smeared at higher values of ΔR, because we allow more events with worse reconstructed top quark to be tagged as top peak.

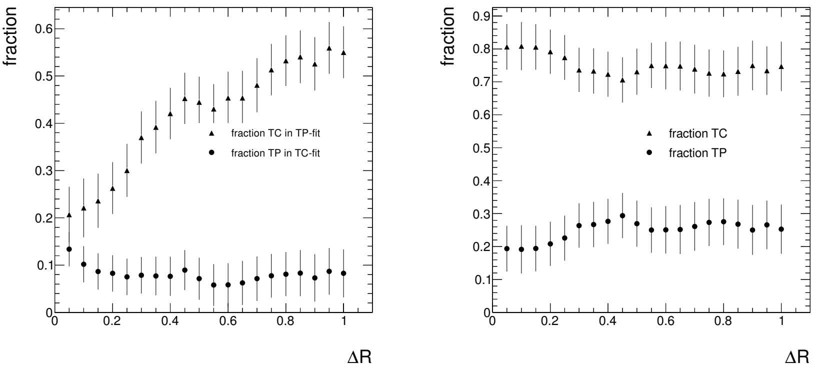
tonic top sample in a peak and combinatorics sample. It is clear already from Figure 6.9 that this splitting will be imperfect. Therefore we have to make sure that our ΔR cut is not biasing our model and also that we can correctly reconstruct both TP and TC contributions in the right proportions. To do this we want to understand the fake rate in our TP-tagged and TC-tagged selection.

If we can show that the fitted yield of correctly reconstructed tops in the TP-tagged and TC-tagged sample combined is independent of ΔR cut, we then have a definition that is independent of simulation truth level parameters, hence it is safe to use it at reconstruction level. The rest of this section is dedicated to this study.

To study the effect of the ΔR cut on the fit we set up a simultaneous fitting procedure. For each ΔR cut value we separate our simulated data into a TP-tagged and a TC-tagged sample. We then fit two models simultaneously to these two samples, where the only difference between the two models is the relative fraction of TP in each MC data sample. We thus assume that the shape of top peak and top combinatorics events are the same in both samples. Our model has five shape parameters:

- top peak described by a Gaussian: $TP-\mu$ and $TP-\sigma$
- top combinatorics described by a TTComb: $TC-\mu$, $TC-\sigma$ and $TC-\alpha$,

and two yield parameters:



(a) (b)

Figure 6.12: Fraction of fitted $TP(TC)$ events in the opposite $TC(TP)$-tagged MC data sample (a) as fitted simultaneously for different ΔR cuts. Total fitted $TP(TC)$ fraction in the complete semileptonic $t\bar{t}$ sample (b) as a function of ΔR. Within the error the fractions stay constant, giving us confidence that our model is correct and choosing a specific cut in ΔR will not bias our results.

- fraction TP in fit to TP-tagged sample: f_{TP}^{TP}
- fraction TP in fit to TC-tagged sample: f_{TP}^{TC}.

Figure 6.10 shows the results of a few simultaneous fits for increasing values of ΔR. Each simultaneous fit is presented by two adjacent plots, the left one showing the TP-tagged MC data in dots and fit result by the solid line while the right one shows the TC-tagged MC data and fit result. For reference each plot contains also a dashed line showing the contribution of the opposite MC data sample, so the left plot shows the fitted TC component of the fit to TP-tagged MC data and in the right one the fitted TP component of the fit to TC-tagged MC data is shown.

For low values of the ΔR cut, the TP-tagged sample is almost purely correctly reconstructed top events and the TC-tagged sample has a substantial contribution from the top peak events. As we increase the ΔR the TP-tagged sample gets more and more contaminated by combinatorics events, while the TC-tagged sample becomes more and more pure combinatorics. For $\Delta R = 0.85$ the fit does not find any TP contribution in the TC-tagged sample.

The question is how our model shape parameters evolve with increasing ΔR. The answer is given by Figure 6.11. The five shape parameters stay constant within the error margins for all values of ΔR. The small increase in the value of $TP - \sigma$ is understood. The TP distribution gets smeared at higher values of ΔR, because we allow more events with worse reconstructed top quarks to slip inside. Some are events where a gluon is radiated off the final state quarks. This gluon is too hard to get correctly accounted for by the jet algorithm, but it is too soft to give the reconstructed jet momentum and jet direction a big disturbance compared to the original quark. Since this gluon still carries away energy, these events have a top mass that is slightly lower giving the TP-tagged distribution a small tail on the low side, that somewhat pulls on the width of the top peak. For a higher statistics sample we expect this tail to be more visible and its effect on the fit more profound.

The evolution of the two yield parameters from the simultaneous fit is shown in Figure 6.12(a). The triangles are the fitted TP yield in the TC tagged sample (f^{TC}_{TP}) as a function of the ΔR cut, while the circles show the fitted TC yield in the TP tagged sample (f^{TP}_{TC}). As expected we see the f^{TC}_{TP} increase with ΔR as we let more events with a worse match contribute to our TP-tagged sample. On the other side with increasing ΔR we see that our TC-tagged sample becomes gradually free of correctly reconstructed top peak events, becoming completely pure combinatorics around $\Delta R \simeq 0.8$.

In Figure 6.12(b) we show the total fitted TP(TC) contribution in the total semileptonic $t\bar{t}$ sample by the circles (triangles). For the total contributions we add the fractions from the TP-tagged sample and the TC-tagged sample as follows:

$$f_{TP} = \alpha \cdot f^{TP}_{TP} + (1-\alpha) \cdot f^{TC}_{TP} \tag{6.9}$$

$$f_{TC} = \alpha \cdot (1 - f^{TP}_{TP}) + (1-\alpha) \cdot (1 - f^{TC}_{TP}), \tag{6.10}$$

where α is the MC data fitted fraction of TP events in the total semileptonic $t\bar{t}$ sample for a specific ΔR cut. What is clear from Figure 6.12 is that although the fitted yields f^{TP}_{TP}, f^{TP}_{TC} evolve with ΔR, the total contribution of $TP(TC)$ is independent of the specific ΔR within error margins.

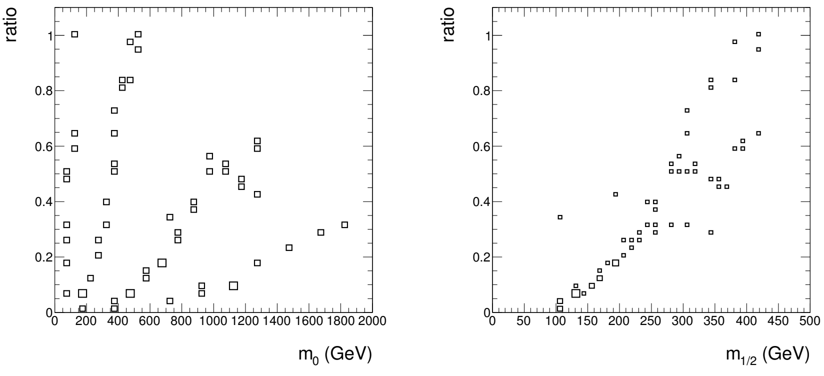
We would like to conclude that although our model has one arbitrary parameter in the ΔR cut, we have shown that it is has no significant influence on the outcome of the fit. Both the shape parameters as well as the total top peak and top combinatorics contributions are constant as a function of ΔR. By choosing a specific value for the ΔR cut, we have a workable definition of our combinatorics and peak samples. The choice for our specific analysis is $\Delta R = 0.1$, which minimizes the fake rate of the TP-tagged sample, as can be seen in Figure 6.12(a).

6.4.4 SUSY Ansatz

The background model will be used to fit the shape parameters of the model in a side-band region (SB) with low $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, and to extrapolate the background yield to a signal-rich region (SIG) at hight $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}. This method assumes that SUSY events have a clear signature of large missing transverse energy due to the LSP escaping detection, and that most background events surviving our primary selection have a characteristic M_{T} of the order of the W mass, due to W bosons decaying into a lepton and a neutrino.

However, if SUSY exists and supersymmetric particles are produced by the LHC, then they are also likely to contribute events to the sideband region. This SUSY contamination has been the subject of much work in the ATLAS collaboration, as it leads to an overestimation of the background in the signal region if unaccounted for. The combined fit method pioneered the contamination assessment procedures as described in [31], by constructing an empirical *Ansatz model* for high mass SUSY models based on the fact that, at least for mSUGRA models, for low $E_{\mathrm{T}}^{\mathrm{miss}}$ and low M_{T} the shapes of the distributions have little dependence with the chosen model point.

This section first describes a way for gross classification of mSUGRA models in terms of SUSY mass scale. We then proceed to answer the question of how to build a 3D model of SUSY contamination that can in principle describe all the different mSUGRA topologies. Once we have built this model, we study the correlation of model shape parameters with the SUSY mass scale for different mSUGRA models. We use these correlations to eliminate redundant



(a) (b)

Figure 6.13: The ratio of the number of SUSY events in the signal region over the number in the control region, as a function of m_0 (a) and $m_{1/2}$ (b) for the SUSY grid points.

degrees of freedom in our model. Finally we compare the fits of this reduced model with the original model to make sure that we still describe all the studied mSUGRA points correctly.

To study different SUSY scenarios and the dependence on model parameters we use the mSUGRA grid as described in section 6.2.2. To be able to grossly classify all these SUSY points in terms of 'SUSY mass', we have defined a new variable:

$$\textbf{ratio} = \frac{N_{SIG}}{N_{SB}}, \tag{6.11}$$

where N_{SIG} (N_{SB}) is the number of events in the signal (sideband) region after selection of section 6.3. Low values of this ratio mean that most SUSY events are in the sideband region , while high values mean that most events are in the signal box. Figure 6.13 shows the distributions of ratio versus the m_0 and $m_{1/2}$ for the grid of SUSY points. The clear correlation between $m_{1/2}$ and ratio confirms that ratio is a good measure of the mSUGRA mass scale. The radial lines in the scalar mass plot mirror the radial lines in the $((m_0, m_{1/2})$-plane) that were used to define the signal grid.

We will now use this ratio to define a new Ansatz model that correctly describes all grid points in the sideband region, while trying to introduce as few new parameters in the combined model as possible.

Figure 6.14 shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the low M_{T} region ($M_{\mathrm{T}} < 150\,\mathrm{GeV}$) for six SUSY grid points sorted by increasing ratio. These points have been selected since they cover a large range in ratio values, and thus a large range of possible mSUGRA signatures. The solid line represents the projection of the SUSY Ansatz model on the $E_{\mathrm{T}}^{\mathrm{miss}}$ axis, for which a TTComb shape was used. What is immediately clear is the gradual evolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$-shape with increasing ratio, the distribution becomes wider and the peak moves to higher values. This shows the possibility of creating an Ansatz model that can describe mSUGRA models in the SB region using the same functional form.

To prove that we see gradual evolution in all the variables figures 6.15 and 6.16 show the M_{T} and M_{jjj} distributions in respectively the low $E_{\mathrm{T}}^{\mathrm{miss}}$ region (SB1 = $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$) and the full L-shaped sideband region. Again the 1D projection of the 3D Ansatz fitted to the

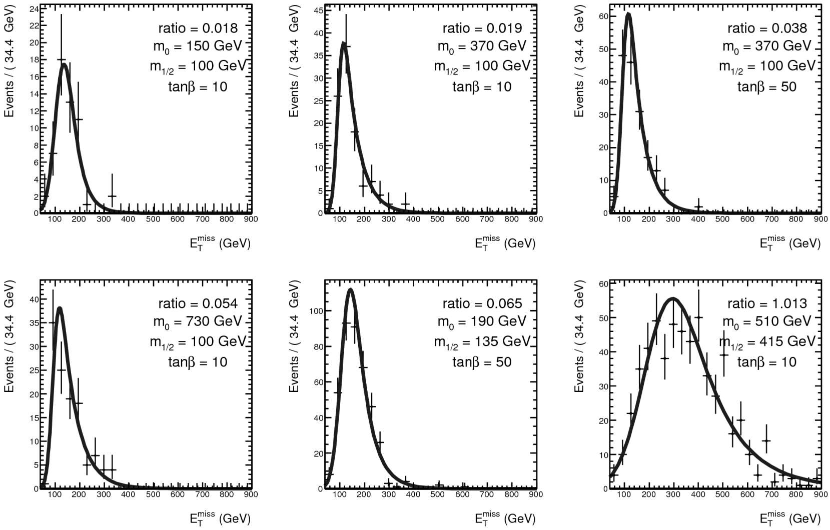


Figure 6.14: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the $M_{\mathrm{T}} < 150\,\mathrm{GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the 1D projection is shown of the 3D model fitted to that point in the SB.

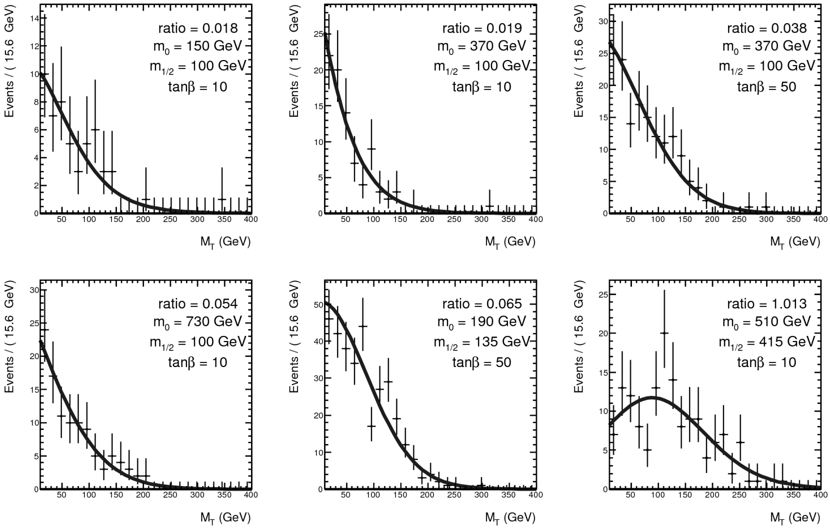


Figure 6.15: M_{T} distributions in the $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the 1D projection is shown of the 3D model fitted to that point in the SB.

MC data for each point is shown with a solid line. The PDF for M_{T} is the TTComb function, while for M_{jjj} it is a convolution of an exponent with a Gaussian. The fits show that the same 3D Ansatz PDF can be used for all our grid points.

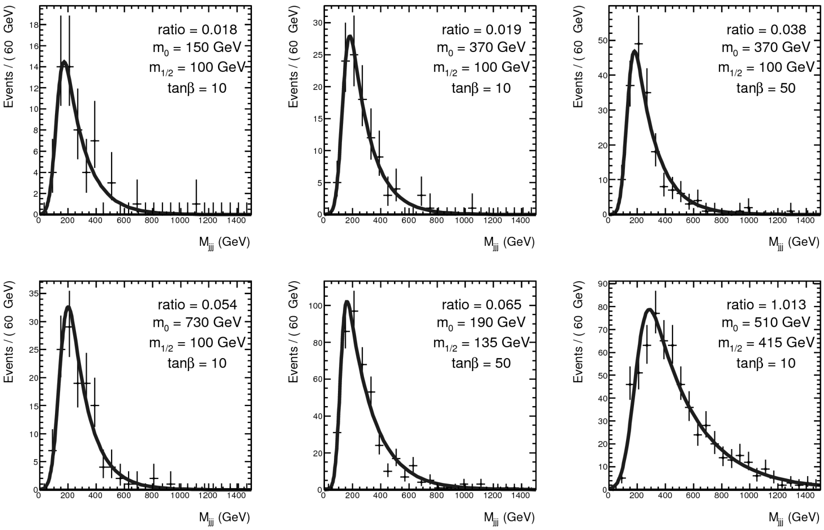


Figure 6.16: $M_{\rm jjj}$ distributions in the $M_{\rm T} < 150\,{\rm GeV}$ and $E_{\rm T}^{\rm miss} < 200\,{\rm GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the 1D projection is shown of the 3D model fitted to that point in the SB.

The great advantage of having a grid of SUSY models instead of just a few benchmark points is that we can study how the parameters of our model vary for the different mSUGRA models. Our initial Ansatz model has nine parameters:

- For missing transverse energy: $E_{\rm T}^{\rm miss} :: \alpha$, $E_{\rm T}^{\rm miss} :: \mu$, $E_{\rm T}^{\rm miss} :: \sigma$

- For transverse mass: $M_{\rm T} :: \alpha$, $M_{\rm T} :: \mu$, $M_{\rm T} :: \sigma$

- For three jet mass: $M_{\rm jjj} :: \tau$, $M_{\rm jjj} :: \mu$, $M_{\rm jjj} :: \sigma$,

where the $M_{\rm jjj} :: \tau = 1/\alpha$ is defined as the inverse of the exponential parameter that was used for the TTComb. One might expect that the shape of the $E_{\rm T}^{\rm miss}$, $M_{\rm T}$ and $M_{\rm jjj}$ distribution changes gradually with increasing mSUGRA mass scale, and hence with increasing ratio. If indeed the parameters are correlated with the ratio, they should also be correlated with each other. As an example, Figure 6.17 compares the values of $M_{\rm jjj} :: \mu$ and $M_{\rm jjj} :: \sigma$ for all grid points. A clear linear dependence can be seen that can be described by a straight line, as shown by the straight line fit in the figure. By replacing the $M_{\rm jjj} :: \sigma$ parameter in our model by a function that describes the linear dependence on $M_{\rm jjj} :: \mu$ the number of parameters is reduced by one.

We consecutively repeat this procedure for the other parameters of our Ansatz model to come to the conclusion that we only need two parameters $E_{\rm T}^{\rm miss} :: \alpha$ and $M_{\rm T} :: \alpha$ to describe all mSUGRA grid models in the SB in three dimensions. The other 7 initial parameters can be described by simple functional relations to the remaining four and to each other as:

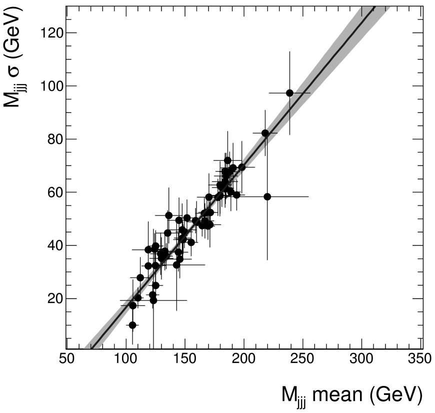


Figure 6.17: Distribution of fitted values of $M_{\rm jjj}$ :: μ and $M_{\rm jjj}$ :: σ for all the grid points. A straight line was fitted to this distribution, with the light grey band indicating the 3 σ error in the fit.

$$
\begin{aligned}
\sigma_{E_{\rm T}^{\rm miss}} &= 7659.78 \times \alpha_{E_{\rm T}^{\rm miss}} + 156.85\,{\rm GeV}, && (6.12)\\
\mu_{E_{\rm T}^{\rm miss}} &= 2.4402 \times \sigma_{E_{\rm T}^{\rm miss}} - 16.459\,{\rm GeV}, && (6.13)\\
\mu_{M_{\rm T}} &= -7815.7 \times \alpha_{M_{\rm T}} - 52.5861\,{\rm GeV}, && (6.14)\\
\sigma_{M_{\rm T}} &= 101.61\,{\rm GeV} \quad (\textit{constant}), && (6.15)\\
\tau_{M_{\rm jjj}} &= 17614.7 \times \alpha_{E_{\rm T}^{\rm miss}} + 436.45\,{\rm GeV}, && (6.16)\\
\mu_{M_{\rm jjj}} &= 0.4339 \times \tau_{M_{\rm jjj}} + 39.567\,{\rm GeV}, && (6.17)\\
\sigma_{M_{\rm jjj}} &= 0.5398 \times \mu_{M_{\rm jjj}} - 37.67\,{\rm GeV}. && (6.18)
\end{aligned}
$$

The two parameter Ansatz model in three dimensions still describes the MC data correctly as shown by Figures 6.18 to 6.20. For comparison the original nine parameter model fit is shown in the same figure by the dotted line, while the light grey band shows the propagated three standard deviations error of this original fit. Note that the slight change in the model shape is small when compared to the statistical uncertainties on the number of events in each bin. At this point we would like to stress that the Ansatz model is only used to account for SUSY contamination in the SB and it makes no assumptions about the shape or abundance of SUSY events in the signal region. Only the SM backgrounds are extrapolated into the signal region, while the SUSY Ansatz only assists in the correct estimation of the backgrounds in the sidebands. This means that although the shape of the SUSY Ansatz does not describe the shape of the mSUGRA distributions perfectly, it will serve well to estimate the amount of SUSY contamination in the SB.

In the next section we will show that this Ansatz indeed estimates the SUSY signal in the control region accurately, and without introducing a significant bias to the method.

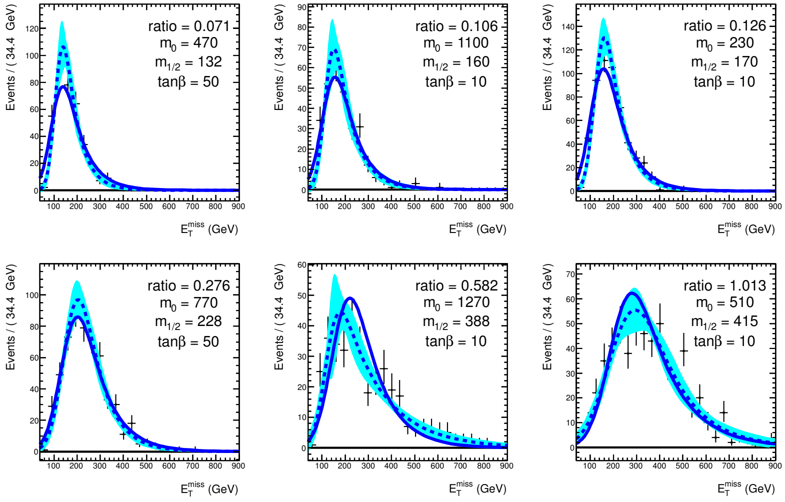


Figure 6.18: $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in $M_{\mathrm{T}} < 150\,\mathrm{GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the solid line is the two parameter Ansatz model fit in the SB and the dotted line is the original nine parameter Ansatz model fit, with a light grey band giving the propagated three standard deviations error of this fit.

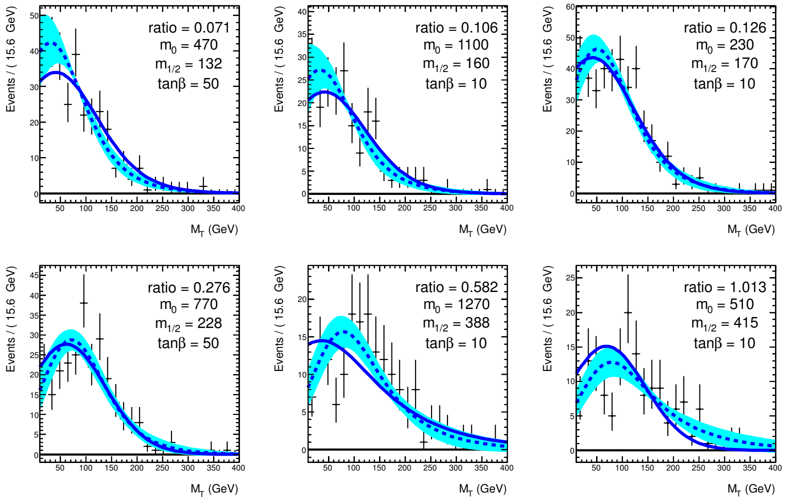


Figure 6.19: M_{T} distributions in $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the solid line is the two parameter Ansatz model fit in the SB and the dotted line is the original nine parameter Ansatz model fit, with a light grey band giving the propagated three standard deviations error of this fit.

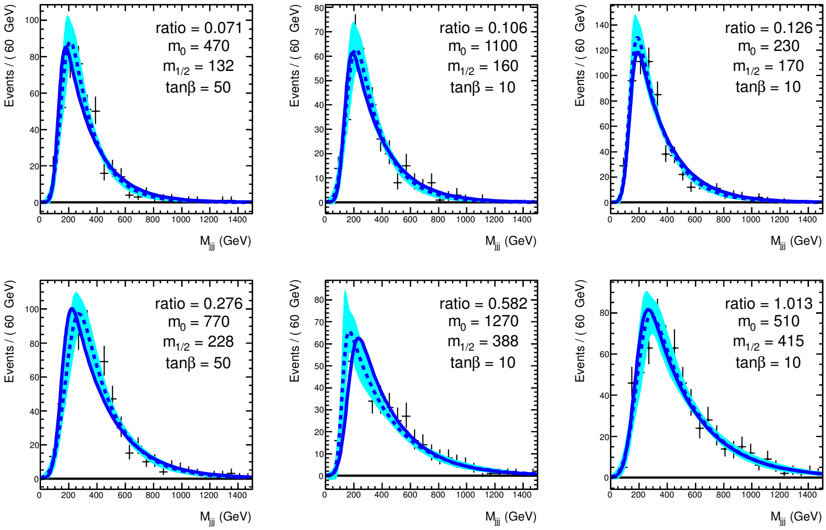


Figure 6.20: M_{jjj} distributions in $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ and $M_{\mathrm{T}} < 150\,\mathrm{GeV}$ region for 6 grid points sorted by increasing ratio. The lowest ratio is in top left, while the highest in the bottom right. For each SUSY point the solid line is the two parameter Ansatz model fit in the SB and the dotted line is the original nine parameter Ansatz model fit, with a light grey band giving the propagated three standard deviations error of this fit.

6.5 Proof of principle

So far we have described the backgrounds we take into account when looking for evidence of supersymmetry, the model that was created to describe these backgrounds, and the model that takes into account the contamination by the SUSY signal leaking into the control region. All that is left to show is that this model accurately estimates the parameters of the combined model in the control region ($E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV}$ or $M_{\mathrm{T}} < 150\,\mathrm{GeV}$) and when extrapolated to the signal region ($E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ and $M_{\mathrm{T}} > 150\,\mathrm{GeV}$) gives an unbiased estimate of the background yields. The boundaries of the control and signal region were based on the shape of the background $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} distributions: They cut behind the W transverse mass peak, and behind the $E_{\mathrm{T}}^{\mathrm{miss}}$ peak of all three backgrounds. In the next chapter, an optimization of the signal region based on the expected significance will be performed.

To show the performance of the method we will present the results of this model in steps of increasing complexity. First we fit the background model to a simulated data sample containing only background, keeping the shape parameters at a fixed value. This fit will be performed both in the FULL range and in the L-shaped control region. The sample we created from MC generated samples of the background, all normalized to 1 fb^{-1}, and combined to form a pseudo dataset.

Then we take into account the fact that the TP/TC separation is not 100% complete, and perform a fit where the TP impurity in the model is fixed to zero. The expected yields from MC will be corrected, and a few shape parameters will be released. The separation of TP and TC events will then be complete in the model.

This fit is used as the baseline for cross checks by performing fits with the complete model on MC data with and without SUSY (with the SUSY samples also normalized to 1 fb^{-1}), and of the background model only to MC data with and without SUSY. We want to make sure that we do not find false positives.

Finally we describe the model with as many shape parameters as possible floating freely in the fit procedure. We check the generator independence of our model by using it to fit MC data containing $t\bar{t}$ samples from different generators. As the last step of the procedure, we also present the result of extrapolation of our combined fit model to the signal region in the presence of SUSY, and show the possible reach into SUSY phase space using this method.

Although we have performed the studies with all the available simulated SUSY samples, for simplicity we demonstrate the behaviour of the method on three *showcase* points. These points were chosen for their different values of *ratio* ($= N_{SIG}/N_{SB}$) and their different SUSY yields in the signal region. The three showcase points have the following parameters:

- Point 1: $m_0 = 1100\,\text{GeV}$, $m_{1/2} = 160\,\text{GeV}$, $A = 0\,\text{GeV}$, $\tan\beta = 10$ and sign(μ)= +, $N_{SIG}/N_{SB} = 208/24$
- Point 2: $m_0 = 77\,\text{GeV}$, $m_{1/2} = 220\,\text{GeV}$, $A = 0\,\text{GeV}$, $\tan\beta = 10$ and sign(μ)= +, $N_{SIG}/N_{SB} = 397/94$
- Point 3: $m_0 = 91\,\text{GeV}$, $m_{1/2} = 300\,\text{GeV}$, $A = 0\,\text{GeV}$, $\tan\beta = 10$ and sign(μ)= +, $N_{SIG}/N_{SB} = 77/45$

The full combined model of SM backgrounds and SUSY contamination that was used for the fits is described in appendix A in terminology of the `RooFit`[46] framework, that was used to perform this study.

6.5.1 Prefit

MC generated samples are used to provide a first estimate for the values of the model parameters. This is done in the *prefit* stage. In this stage the TP model is fitted to a MC sample with only TP events, the CW model is fitted to a CW sample, and $T2$ is fitted to a sample of $T2$ events. These three fits are done simultaneously, such that the global parameter values are determined using information from all relevant background samples. Parameters that are specific for one background model are fitted only to that sample. The SUSY model is fitted separately, since it has no parameters in common with the backgrounds model, and furthermore is only defined in the L-shaped control region.

The result of the SM backgrounds prefit is shown in figure 6.21. The SUSY prefit on showcase point 2 is shown in figure 6.22, separately for each leg of the L-shape. Note that the TP model still takes some impurity by top combinatorics events into account.

6.5.2 Combined background fit with shape parameters fixed

After getting an estimate of all the shape parameters, we add the three background models to form one combined model as discussed in section 6.4.1. The first test of this model is to leave all shape parameters constant and float only the yields of the three background contributions. For the data points we use a set of simulated background samples for the three backgrounds, all normalized to 1 fb^{-1}. Since we use the same samples to perform the prefit, the fitted yields should be comparable to the true yields if our model is anywhere near correct.

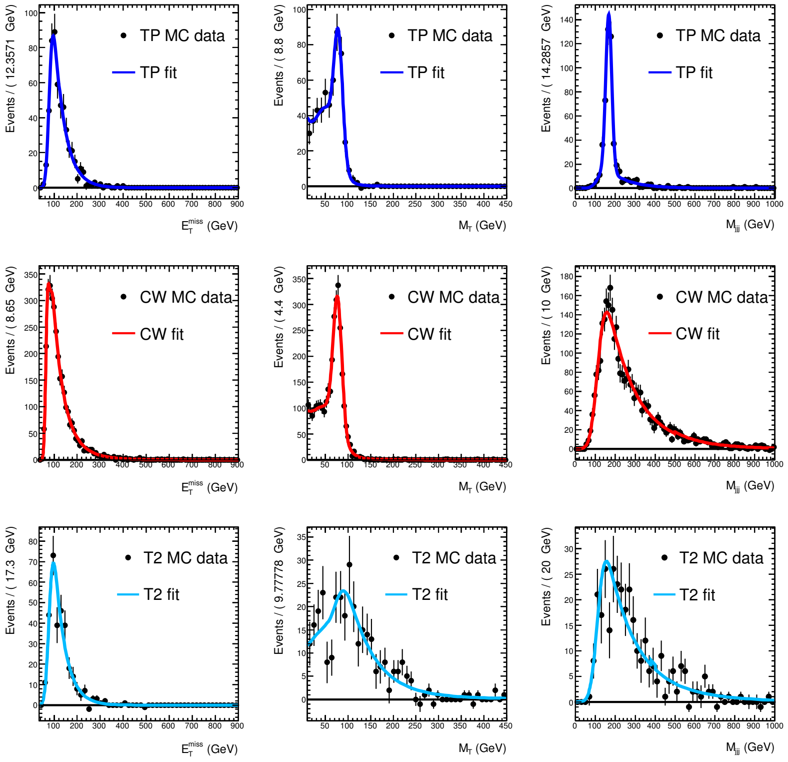


Figure 6.21: The prefit result for the backgrounds. For each background this figure shows three 1D projections of the 3D pdf. The three 3D models are fitted simultaneously to their respective background sample. The small excess in the CW M_{jjj} distribution is due to TP contamination that was not accounted for in the model.

The results of the simplest combined fit are shown in Table 6.9. Here the yields as obtained from the fit are compared to the true number of events in each sample. The fit is done twice: once in the full phase space region $40 < E_{\mathrm{T}}^{\mathrm{miss}} < 900\,\mathrm{GeV}$ and $10 < M_{\mathrm{T}} < 450\,\mathrm{GeV}$, and once in the L-shaped control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV} \vee M_{\mathrm{T}} < 150\,\mathrm{GeV}$. As can be seen from this table, the fitted yields are correct within approximately one standard deviation. Since we define our model using recursive fractions as in equation 6.2, the sum of the fitted yields equals the true MC yields by construction. In the next sections all fits are done in the L-shaped control region.

6.5.3 Fit with a pure TP model: minimal floating shapes

The splitting of the semileptonic $t\bar{t}$ sample, as was discussed in section 6.4.3, is not perfect. Some combinatorics (TC) events are still present in the peak (TP) sample and vice versa. In order to get a first estimate of the shape parameters in the prefit, this contamination is taken into account in the TP M_{jjj} model. However when performing the combined fit, we expect these events to contribute to the W+combinatorics (CW) fraction, and we manually set this

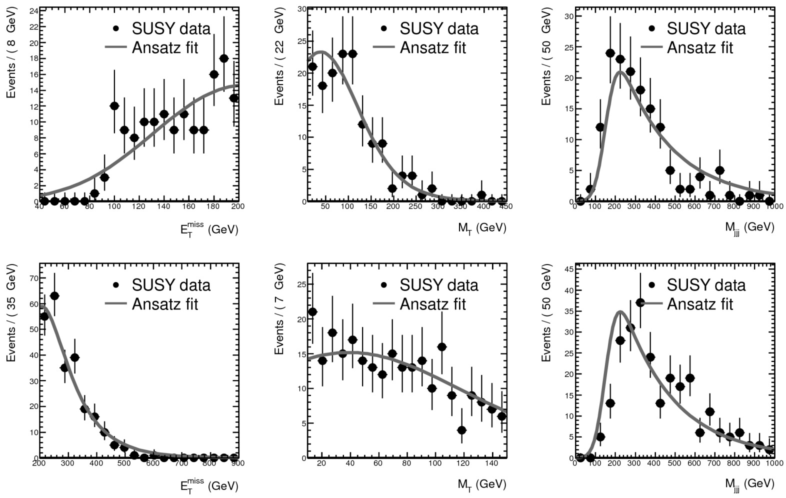


Figure 6.22: The prefit result for the SUSY showcase point 2. The three top plots are for the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ leg of the L-shaped region ($E_{\mathrm{T}}^{\mathrm{miss}} < 200\mathrm{GeV}$), while the bottom plots are for the low-M_{T} leg ($M_{\mathrm{T}} < 150\mathrm{GeV}$, $E_{\mathrm{T}}^{\mathrm{miss}} > 200\mathrm{GeV}$).

TP contamination fraction to zero in the model, thus we fit only the size of the hadronic top mass peak, without contamination. If we want to do a first test of the combined model with a contamination-free TP model, we need to float the M_{jjj} parameters of the CW and TP models, since these are expected to change with respect to the prefit. This we call the combined fit with *minimal floating shapes*, of which the results are shown in Table 6.10. This test is but a slight extension on the fixed shapes fit of the previous section.

In Table 6.10 a correction has been applied to the true yield of the background samples. By setting the peak fraction of the TP model to 1, the fitted TP yield will only be sensitive to the top mass peak, and not to the combinatorics contamination that ended up in the TP sample due to imperfect splitting. Thus the fitted TP yield will be different from the number of events in the TP sample. The applied correction makes sure that we can still have a meaningful comparison between fitted and true yield.

From the study in section 6.4.3 we have an estimate of the true number of correctly reconstructed top events and combinatorics background events in the semileptonic $t\bar{t}$ sample. We use this to compare our fit result to our estimated number of TP and TC using $N_{est}^{TP/TC} = N_{true}^{T1} * f_{TP/TC}$, with N_{true}^{T1} the total number of events in our semileptonic $t\bar{t}$ sample, $f_{TP} = 0.1915$ the estimate of the TP fraction in the semileptonic $t\bar{t}$ (T1) sample, and $f_{TC} = 0.8085$ the estimated TC fraction. As we showed in section 6.4.3 there is a small uncertainty on the true TP/TC yield related to the cut on ΔR, however we have shown that this error is small enough to be ignored.

Thus the minimal floating shapes fit allows us to complete the splitting of TP and TC , and to compare our fit result reliably to the true yields. Table 6.10 shows that also with minimal floating shapes the combined fit performs well.

Sample name	Fitted Yield	True Yield
FULL region		
TP	566 ± 51	514
$T2$	320 ± 32	311
CW	3046 ± 60	3107
Control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV} \vee M_{\mathrm{T}} < 150\,\mathrm{GeV}$		
TP	565 ± 51	514
$T2$	288 ± 32	299
CW	3065 ± 59	3105

Table 6.9: Fitted yields estimated by fitting the combined model with fixed shapes in the full ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) region and in the L-shaped control region to a background only sample.

Sample name	Fitted Yield	True Yield
Control region $E_{\mathrm{T}}^{\mathrm{miss}} < 200\,\mathrm{GeV} \vee M_{\mathrm{T}} < 150\,\mathrm{GeV}$		
TP	412 ± 54	400
$T2$	293 ± 31	299
CW	3213 ± 61	3219

Table 6.10: The results of the fit in the L-shaped control region with TP and CW M_{jjj} shapes floating, also called the minimal floating shapes fit in the text..

6.5.4 Bias check

A cross check of the validity of the combined fit method is to study the behaviour of the model with SUSY component fitted to MC data without a SUSY component and vice versa. The results of these cross checks are listed in Table 6.11. Here the model was fitted with minimal floating background shapes and the two SUSY parameters floating. For completeness the table also shows the fit results with no SUSY component in either model or MC data, and with SUSY component in both the model and the MC data. The model with SUSY component fits a SUSY fraction of 0 when there is no SUSY in the MC data, as one would expect. We expect a bias in the yields only when fitting a model without SUSY component to a sample containing SUSY events. Indeed the table shows that in this case the $T2$ component gets overestimated, since that is the only background with a high $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} tail somewhat like SUSY models.

6.5.5 The combined fit and extrapolation into the signal region

The goal of the combined fit method is to estimate the background contribution and shapes in an L-shaped control region and extrapolate these to the signal region. There we perform a counting experiment comparing the number of observed events with the number of expected background events. Table 6.12 shows the result of such a fit with minimal floating shapes*and* extrapolation for the background model fitted to a background-only sample. We see that the extrapolation yields very accurate results for the estimated background contribution, showing

Sample name	Fitted Yield	True Yield
Fit with no SUSY in MC data, no SUSY in pdf		
TP	412 ± 54	400
$T2$	293 ± 31	299
CW	3213 ± 61	3219
Fit with no SUSY in MC data, SUSY in pdf		
TP	411 ± 54	400
$T2$	293 ± 31	299
CW	3213 ± 61	3219
SU	0 ± 8	0
Fit with SUSY in MC data (showcase point 1), SUSY in pdf		
TP	383 ± 57	400
$T2$	309 ± 48	299
CW	3278 ± 67	3219
SU	157 ± 48	208
Fit with SUSY in MC data (showcase point 2), SUSY in pdf		
TP	379 ± 53	400
$T2$	321 ± 41	299
CW	3255 ± 67	3219
SU	360 ± 39	397
Fit with SUSY in MC data (showcase point 3), SUSY in pdf		
TP	405 ± 53	400
$T2$	318 ± 34	299
CW	3212 ± 63	3219
SU	61 ± 17	77
Fit with SUSY in MC data, no SUSY in pdf		
TP	326 ± 52	400
$T2$	455 ± 35	299
CW	3137 ± 61	3219

Table 6.11: The cross check of the combined fit, fitting with and without a SUSY model in the pdf, with and without SUSY in our MC data sample. For the background model a pure TP sample was used, the SUSY shape parameters were floating as described in section 6.4.4.

Sample name	Fitted Yield in SR	True Yield in SR
TP	0.068 ± 0.009	0
$T2$	11.1 ± 1.2	12
CW	1.54 ± 0.03	2

Table 6.12: The results of the fit in the control region with TP and CW M_{jjj} shapes floating, with the yields extrapolated to the signal region (SR).

that our model behaves well both in the control and in the signal regions.

6.5.6 The combined fit with floating shape parameters

To reduce dependence on any single MC generator as much as possible, we want to release and float as many shape parameters as possible. The full model has too many parameters to float them all, but we float as many as possible while keeping the fit stable. Determining which parameters can be released in the fit requires some careful handling. To determine what set of floating shapes is acceptable, we looked at the change induced by floating shape parameters in the estimated number of background events in the signal region. The result of the fit with only the TP parameters floating, which we expect to yield the best possible result, was used as the baseline. The model was fitted to a MC data sample containing mSUGRA showcase point 2.

We first floated all parameters of the fit one by one, determining which parameter induced the smallest error when released in the fit. We then ordered the parameters according to this error. To avoid bias we did not want to introduce flexibility in one component while keeping the shape of another fixed, so we then grouped the parameters such that there was a parameter of each model in each group. Then we started added groups of parameters, and evaluated the outcome of the fit. The relative change in the extrapolated SM yield was below 10%, as can be seen in Table 6.13. The final set of parameters floated in the maximum floating shapes fit is quoted in appendix A. When this list was expanded further, two parameters of the $T2$ model destabilized the fit, such that the minimization procedure failed to converge. The group with only CW parameters was not added in order to avoid the aforementioned bias.

We will quote the full result here for the three showcase points on the mSUGRA grid. For this fit, we float the largest subset of shape parameters possible. This includes the global M_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ parameters, most CW and TP parameters and most $T2$ M_{jjj} and $E_{\mathrm{T}}^{\mathrm{miss}}$ parameters. The $T2$ M_{T} parameters are all set to the values estimated from the simulated sample. The fit gives us three results: the relative background to signal contribution in the control region, the relative contribution of the different processes to the total SM background, and through extrapolation it gives the total number of expected SM events in the signal region. The respective yields are listed in Table 6.14 both in the control (CR) and in the signal region (SR). Also quoted is the significance of detecting the signal over the expected background for the observed number of events. Significance was calculated using the tools provided by the `RooStats` framework [62, 63]. The problem is treated in a fully frequentist fashion by interpreting the relative background uncertainty as being due to the auxiliary sideband observation, or fit result in our method, while the number of observed events and the number of expected background events are distributed as Poissons, neglecting systematic errors.

group	$\Delta N_{SM}^{exp}(\%)$
`G_mtrans_sigma_ratio`, `T2_mtrans_sigma_ratio`	3%
`G_mtrans_corefrac`, `T2_mtrans_corefrac_j4`	2%
`CW_mtrans_base`, `TP_mtrans_base`, `G_mtrans_sigma`	6%
`CW_etmiss_mean`, `G_etmiss_mean`	2%
`T2_mjjj_base_sig`, `TP_mjjj_base`, `T2_mjjj_mean_slope`	2%
`CW_etmiss_sigma`, `T2_etmiss_sigma`, `TP_etmiss_sigma`	*not converged*
`CW_etmiss_base`, `T2_etmiss_base`, `TP_etmiss_base`	*not converged*
`CW_etmiss_base_ratio`, `CW_etmiss_frac`, `CW_etmiss_mean_ratio`	*Not balanced*

Table 6.13: The groups of parameters that were released together in the fits with floating shape parameters. The naming of the parameters is as follows: `CW_etmiss_mean` denotes the mean of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb of the *CW* model. See for a detailed description of the model appendix A. The first 5 parameter groups together with the *TP* M_{jjj} and *SU* Ansatz parameters form the complete set of floating shape parameters. If the width and base of the $E_{\mathrm{T}}^{\mathrm{miss}}$ PDFs was floated in the fit, the minimization procedure failed to find convergence. The *CW* $E_{\mathrm{T}}^{\mathrm{miss}}$ parameters were not floated to avoid introducing bias due to more freedom in the fitting of one model.

The last result, the extrapolated SM background model into the signal region is shown in figure 6.23. It shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T} distribution of the MC data sample (filled circles) and the SM (squares) and SUSY (open circles) MC data separately, together with the background model extrapolated to the signal region at $M_{\mathrm{T}} > 150\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$.

6.5.7 Significance reach

Figure 6.24 shows the significance of the excess found using this method for all points in the SUSY phase space for which we have a simulated sample. The points that are expected to be found with high significance are those points which have a low enough mass scale to be copiously produced at the LHC, but not so low that all signal events end up in our control region. This can be clearly seen in figure 6.24 as for low m_0 and $m_{1/2}$ the significance is below one, steadily rising with increasing scalar and fermion masses well above the discovery potential of significance equal to 5. For higher values of m_0 and $m_{1/2}$ however, the number of signal events in the signal region drops down to levels that cannot be measured with enough significance at the chosen integrated luminosity. Note that the significance in these plots are calculated with no systematic error.

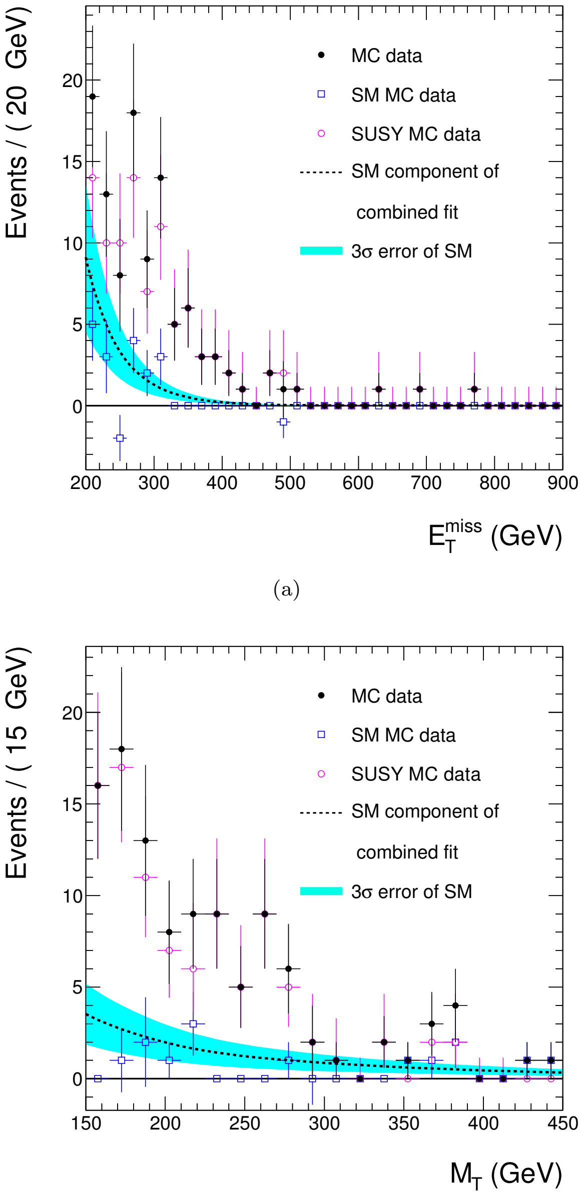
Figure 6.25 shows the comparison of the measured cross section used in making the significance plot 6.24 compared to the true cross section for each point. This plot shows a small positive bias, which means that the significances in figure 6.24 are slightly overestimated. This is caused by a systematic, but small $\mathcal{O}$(1 event) underestimation of the background yields. For those parts of SUSY phase space where a significant measurement can be made, this bias is much smaller than the poisson error on the number of events in the signal region.

6.5.8 Generator independence

To test the model with floating shapes, we did the prefit on the MC data sample with $t\bar{t}$ events generated using MC@NLO, and then constructed a MC data sample with $t\bar{t}$ events generated using ALPGEN. If our model is indeed generator independent, it should be able to correctly

Sample name	CR Fitted Yield	CR True Yield	SR Fitted Yield	SR True Yield
		Showcase point 1		
TP	428 ± 61	400	0.03 ± 0.09	0
$T2$	406 ± 68	299	13 ± 2	12
CW	3113 ± 87	3219	0.9 ± 0.5	2
SU	179 ± 52	208	24± 2.	24
			Significance = 4.3	
		Showcase point 2		
TP	395 ± 56	400	0.02 ± 0.07	0
$T2$	365 ± 64	299	12 ± 2	12
CW	3174 ± 88	3219	1.2 ± 0.6	2
SU	380 ± 42	397	95± 2	94
			Significance = 13	
		Showcase point 3		
TP	448 ± 61	400	0.03 ± 0.08	0
$T2$	392 ± 55	299	13 ± 1.79	12
CW	3088 ± 86	3219	1.1 ± 0.6	2
SU	67 ± 19	77	45± 2	45
			Significance = 7.7	

Table 6.14: The results of the combined fit in the sidebands, with the largest possible subset of parameters floating. Three different MC datasets were created using the three showcase points. The estimated and true yields are shown for the control region (CR) as well as extrapolated to the signal region (SR). The errors quoted are the statistical errors on the extrapolation from the fit. Note that the Poisson error on the event count in the signal region is not included.



(a)

(b)

Figure 6.23: These two figures show the extrapolation to the signal region of the background component of the model after fitting the complete model in the L-shaped region, with 3σ error band in both E_T^{miss} (a) and M_T (b). As reference the respective SM and SUSY MC data is shown.

Sample name	CR Fitted Yield	CR True Yield	SR Fitted Yield	SR True Yield
MC@NLO prefit model, ALPGEN $t\bar{t}$ in MC data sample				
TP	314 ± 56	255	0.00 ± 0.01	0
$T2$	314 ± 45	328	8 ± 4	10
CW	3031 ± 79	3046	0.03 ± 0.06	2
SU	367 ± 62	397	98± 4	94
			Significance = 7.7	
ALPGEN prefit model, MC@NLO $t\bar{t}$ in MC data sample				
TP	423 ± 60	400	0.00 ± 0.02	0
$T2$	367 ± 57	299	10 ± 6	12
CW	3214 ± 94	3219	1.6 ± 0.9	2
SU	311 ± 58	397	96± 6	94
			Significance = 6.0	

Table 6.15: The results of the combined fit tested on MC data with a different $t\bar{t}$ generator. This MC data sample was created using showcase point 2. The top table shows the result of a model prefitted on MC@NLO $t\bar{t}$, and then fitted on a MC dataset containing ALPGEN $t\bar{t}$, the bottom table contains results for the reversed test. For the ALPGEN samples, the analysis of the TP/TC yield in semileptonic $t\bar{t}$ was redone to obtain an estimate of the true number of TP and TC events. Note that the total number of ALPGEN top events is lower than that for MC@NLO. This is due to an incorrect k-factor that was used when making this table.

estimate the yield of ALPGEN $t\bar{t}$ events in this new MC dataset. Table 6.15 shows that indeed the yields of $t\bar{t}$ are correctly estimated. Unfortunately we did not have an alternative W+jets generator available, but through TP/TC splitting and W/TC recombination at least changing the $t\bar{t}$ generator affects all background models. The same table shows the reverse test, a model prefitted with the ALPGEN sample fitted to a MC dataset with MC@NLO $t\bar{t}$. This result shows a slight underestimation of SUSY, due to an overestimation of dileptonic $t\bar{t}$. This means that it is worthwhile to look for ways to further stabilize the fitting procedure, enabling the release of some shape parameters of the dileptonic $t\bar{t}$ model. In the conclusion of this thesis a couple of methods will be discussed to achieve this. For now, we include the generator dependence as a systematic error. This systematic error is not very large since the $T2$ overestimation does not influence the final estimate of the number of SUSY events in the signal region too much.

6.5.9 Validation of the fit procedure

To verify the absence of bias in our fitting procedure and to study the correctness of the quoted errors we have run a series of 'toy' Monte Carlo studies on different SUSY grid points. For each shown SUSY grid point we ran 1000 experiments, where we sample the fitted shape of the combined pdf to generate simulated events and perform a fit on the generated data. Figure 6.26 shows the results of 1000 experiments for showcase point 2, where we use the model with minimal floating shapes[3] as described in section 6.5.3. The fraction of SUSY events, f_{SU}, and

---

[3] For the lack of available computing power.

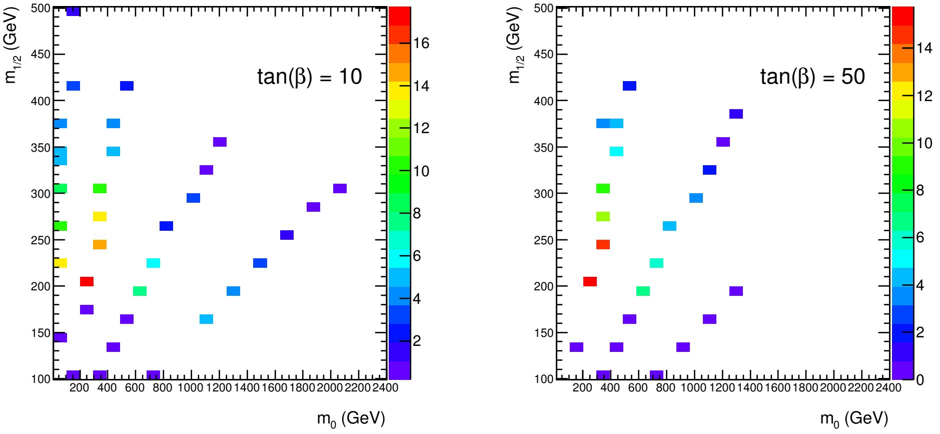


Figure 6.24: Significance that can be obtained with the combined fit method as a function of m_0 and $m_{1/2}$, for $\tan\beta = 10, 50$.

the corresponding error that come out of the fit are used to calculate the pull, defined as:

$$\text{pull} f_{SU} = \frac{f_{SU}^{fit} - f_{SU}^{true}}{\sigma_{f_{SU}^{fit}}}. \tag{6.19}$$

The pull distribution can be described by a Gaussian function, that should have a mean equal to 0 if there is no bias and the width of the pull distribution should be 1 if the error estimate is correct. This follows from the properties of maximum likelihood estimators, that ensure that in the limit of an infinite amount of MC data, the estimates are normally distributed around the true value[64]. However in the scenario under study, where the yield of SUSY is quite small (order of $1 - 10\%$ of the total number of background events in control region), it is typically somewhat overestimated[64, 65]. The yield value and error distributions become asymmetric, as shown for f_{SU} and $\sigma_{f_{SU}}$ in Figure 6.26, as a result of trying to keep the yield physical, hence above zero. As the yield approaches 0, the low error becomes smaller then the high error, leading to the probability of an upward fluctuation becoming greater then the probability of a downward one, giving rise to a small overestimation of the yield and an asymmetric pull distribution.

As can be seen from Figure 6.26, there is indeed a small bias in our procedure. As the asymmetric errors tend to give pull values in the tails of the pull distribution, it is worthwhile to perform a *central fit* to the core of the pull$[-2, 2]$ distribution. Both the central fit and full (phase space) fit results are shown in Figure 6.26 as respectively a solid and a dotted line as well as the fit results. For this SUSY point the central fit shows us an unbiased and correctly estimated pull, while the full fit has a σ significantly away from 1.

To make sure that it is not only our specific showcase point 2 that is somewhat overestimated by the toy experiments, in Figure 6.27 we show the pull distributions for 1000 toy experiments for the other two showcase points as well. There seems to be a small upward bias in the estimation of the SUSY yield, even if we perform a central fit on the pull distributions.

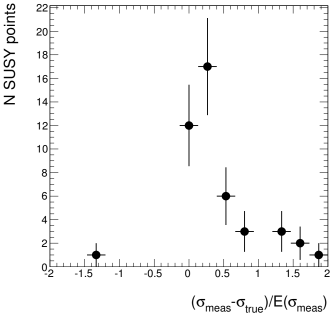


Figure 6.25: Bias check on the fit results that were used for figure 6.24. This plot shows the difference in the SUSY cross section as measured from the fit and the true SUSY cross section, divided by the error on the measurement. A small positive bias is visible, indicating that this method tends to overestimate the SUSY yield. Only those points on the mSUGRA grid are displayed here that yield more than 1 event in the signal region at 1 fb^{-1}.

We can however state that our procedure estimates the errors correctly as the width of the pull distributions is equal to 1 within error margins.

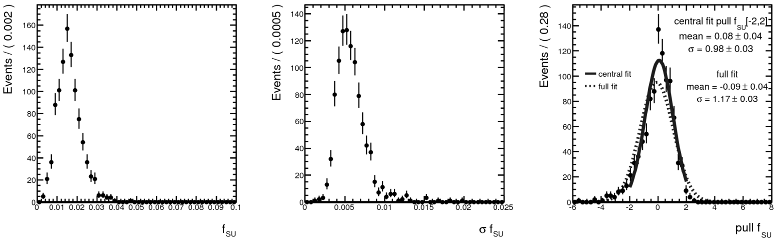


Figure 6.26: The fitted fraction (left) of SUSY, f_{SU}, the error (middle) on the fitted SUSY fraction, σf_{SU} and the pull distribution (right) of f_{SU} for the showcase point 2. A Gaussian function fit was performed to the full pull distribution (solid line) and to its central part $[-2, 2]$ (dashed line). The found *mean* and σ are quoted with errors for both fits.

6.5.10 Systematics

We want to find a measure of what influence detector effects like misreconstruction of jets, leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ have on the power of our analysis. We varied the parameters listed in Table 6.16 and rerun the combined fit on a MC dataset with background plus the MC data for mSUGRA showcase point 2, where the parameter variation was taken as in [31]. The variation quoted here is the variation of the measured SUSY cross section. The cross section is defined as:

$$\sigma_{SUSY} = \frac{N_{obs} - N_{exp}}{\epsilon_{sel}}. \tag{6.20}$$

The uncertainties on the systematic error that are listed in the table are due to Monte Carlo statistics. These uncertainties may occasionally be dominated by small event counts (e.g. the

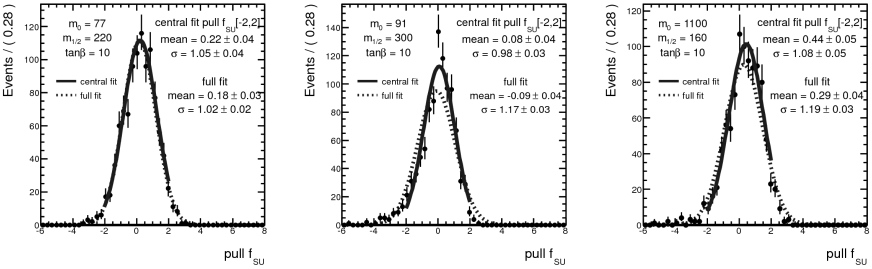


Figure 6.27: The pull distributions of f_{SU} for the three showcase points. The values of both the central and the full fits are quoted, while the resulting functions are shown by respectively solid and dashed lines.

number of events that migrate due to the systematic check), and may limit the precision of these systematic checks. To estimate the MC generator dependence we used the test described in the previous section which shows the largest deviation with the ALPGEN prefitted model tested on a MC data containing MC@NLO $t\bar{t}$.

Variations in the jet energy scale and resolution appear to have the largest effect on the outcome of the method, giving a systematic uncertainty around 10 %. The other systematic uncertainties are at the level of a few percent.

Finally there is a systematic error associated with the parameters of the model that were not fitted, but kept at the value determined in the prefit stage. The size of this systematic error is correlated to the size of the systematic error of the MC generator, since the shape of the distributions are different when a different generator is used. A more detailed study on the effect of keeping certain shape parameters fixed is performed in the next chapter, when the combined fit method is applied to data.

Source	Systematic Error [%]
Jet Energy scale up 5%	11 ± 2.4
Jet Energy scale down 5%	7.2 ± 2.6
Jet Energy resolution 10%	12 ± 2.9
Muon Energy scale up 0.2%	1.2 ± 0.8
Muon Energy scale down 0.2%	0.72 ± 0.46
Muon Energy resolution 1%	1.4 ± 0.94
Soft Missing ET scale up 10%	2.6 ± 1.1
Soft Missing ET scale down 10%	1.9 ± 1.5
MC generator dependence	0.32 ± 0.58

Table 6.16: Different sources of systematic error and their effect on the SUSY combined fit performance.

Chapter 7

Combined Fit on data from 7 TeV pp collisions

7.1 Introduction

In the previous chapter the combined fit method was developed, and a model was designed to describe the data simulated for *pp* collisions at 10 TeV. In this chapter the same method will be applied to data measured by the ATLAS detector from 7 TeV collisions delivered by the LHC in the year 2010.

The changed center of mass energy together with the smaller size of the sample ($35\,\text{pb}^{-1} \pm 3.4\%$ [66] compared to the $1\,\text{fb}^{-1}$ used in the previous chapter) will lead to an adaptation of the model developed in the previous chapter. This chapter will discuss the online and offline selection of the measured data, the new simulated samples, and the adapted combined fit model. The definition of objects and event selection criteria will be adapted, as will the definition of the control and signal regions. New MC simulations will be used to define and test this model. It will be shown that the correlations between observables are small enough relative to the statistical uncertainty to be ignored for this relatively small dataset. Furthermore a new SUSY Ansatz model will be developed that describes the expected behavior of the signal in a region of mSUGRA phase space that can be tested using the 2010 dataset.

After validation of the new model on simulated data, a fit of the model to the measured data will yield an expectation of the SM background in the signal region. Using a profile likelihood method this expectation can be compared to the observed number of events, resulting in either a discovery of an excess of events over the SM expectation, or a new physics model independent limit on the amount of signal events present in data. The excess or limit that is found will then be compared to the simulated mSUGRA models.

7.2 Data selection

7.2.1 Triggering

Online selection of interesting events was done by requiring a muon trigger in each event. The specific trigger used changed over the period of data taking, as trigger requirements shifted with

Period	Trigger	$\int Ldt$ (pb^{-1})
D	L1_MU6	0.288
E	EF_mu10_MSonly	0.937
F-H	EF_mu13	14.4
I	EF_mu13_tight	19.0

Table 7.1: The triggers used for the different data periods. The number in the trigger name denotes the muon p_T threshold.

increasing instantaneous luminosity delivered by the LHC. Starting with only the L1 trigger, later the L2 and EF triggers were added to further reduce the amount of data recorded by ATLAS. To keep the trigger rate under control, the trigger thresholds were gradually increased.

The addition of the high level triggers, and the increase of the trigger threshold does not have a significant effect on the number of events that passed the selection. The offline event selection asks for one muon with a $p_T > 20\,\mathrm{GeV}$. The L1 efficiency of each trigger for finding such a muon is the same for a wide range of trigger thresholds. This can be seen from the shape of the muon trigger turn on curve, a steep rise around the threshold followed by a plateau at higher p_T. See Figure 4.11(a) in section 4.5, which shows that the efficiency of the muon trigger is about 75% for the L1 muon trigger with a p_T threshold of 6 GeV. Furthermore for an offline muon with $p_T > 20\,\mathrm{GeV}$ the L2 and EF triggers are close to 100% efficient with respect to the L1 trigger, see Figures 4.11(b) and 4.12.

Data taking was divided in different periods, labeled A-I, when changes to the LHC or the detector were made. The data used in this chapter was taken during period D-I. Periods A-C contained only 16 nb^{-1}. Table 7.1 shows the triggers for the different periods of data taking, and the amount of integrated luminosity recorded with that trigger. For MC an EF trigger requirement with a p_T threshold of 10 GeV was used. Again, as the offline lepton selection asks for muons with a p_T of at least 20 GeV, well into the plateau region of both the data and MC triggers, the efficiencies of both triggers is compatible.

7.2.2 Run selection, Integrated luminosity

After passing the trigger, events are recorded and reconstructed. Not all recorded events contain data of sufficient quality to be suitable for physics analysis. A subdetector malfunction, or issues with beam stability could have effected the recorded data. Therefore recorded events went through a first selection requiring stable beams from the LHC, and all detector subsystems in operation. Besides an overall flag set by the ATLAS run coordinator each subsystem and trigger system records a flag for each run indicating smooth running of the system, and absence of abnormalities in the recorded data., e.g. symmetric ϕ distribution within acceptance and expected occupancy in different regions of the subdetector. These selection requirements resulted in a total integrated luminosity of 35 pb^{-1} ± 3.4% [66].

During period D, with a total integrated luminosity of 288 nb^{-1} the RPC reported ongoing issues, and data was accepted with less than 90% of the RPC functioning. The problematic RPC data was then reprocessed offline in order to be usable for physics analysis.

Process	Cross Section (pb)
W+jets + l + 0 partons	$8.5 \cdot 10^3$
W+jets + l + 1 partons	$1.6 \cdot 10^3$
W+jets + l + 2 partons	$4.5 \cdot 10^2$
W+jets + l + 3 partons	$1.2 \cdot 10^2$
W+jets + l + 4 partons	31
W+jets + l + 5 partons	8.5

Table 7.2: The cross section for the W + jets + 1 lepton ALPGEN samples, with NNLO k-factor applied. These samples were generated for all three lepton flavors.

7.2.3 MC samples

In the previous chapter the generators and simulation packages used to create MC samples for signal and background were described in some detail. The MC samples used in this analysis were mostly created using the same tools. This means $t\bar{t}$ events were generated using MC@NLO and W+jets events using ALPGEN. ALPGEN was also used to create a second sample of $t\bar{t}$ events, used to estimate the systematic effect of using a specific generator. The different center of mass energy used in these samples resulted in different cross sections compared to the ones used in the previous chapter. For $t\bar{t}$ generated with MC@NLO, a NNLO cross section of 159 pb was used, with the generator level filter efficiency (selecting only leptonic $t\bar{t}$ decays) of 0.56. For ALPGEN $t\bar{t}$, the cross sections are listed in Table 7.3. For W+jets, the cross sections are listed in Table 7.2.

Process	Cross Section (pb)
$t\bar{t} \to l\nu l\nu$ + 0 partons	5.8
$t\bar{t} \to l\nu l\nu$ + 1 partons	5.7
$t\bar{t} \to l\nu l\nu$ + 2 partons	3.6
$t\bar{t} \to l\nu l\nu$ + 3 partons	2.3
$t\bar{t} \to l\nu qq$ + 0 partons	24
$t\bar{t} \to l\nu qq$ + 1 partons	24
$t\bar{t} \to l\nu qq$ + 2 partons	15
$t\bar{t} \to l\nu qq$ + 3 partons	9.4

Table 7.3: The cross section for the $t\bar{t}$ ALPGEN samples, with an NLO k-factor applied. Samples were generated for all three lepton flavors.

For QCD specific samples were generated using PYTHIA, where one muon was already requested at generator level, thus reducing the number of events that need to be generated and processed for this analysis. The cross sections for these QCD samples are listed in Table 7.4.

A new grid was defined in the (m_0,$m_{1/2}$) plane of mSUGRA phase space, with $A_0 = 0\,\text{GeV}$, $\tan\beta = 3$ and $sgn(\mu) = +$, and with ($80 < m_0 < 850\,\text{GeV}$) and ($100 < m_{1/2} < 340\,\text{GeV}$). The spacing of the grid was twice as close for lower values ($m_0 < 440\,\text{GeV}$ and $m_{1/2} < 220\,\text{GeV}$), the region where exclusion/discovery was to be expected. For each point of the grid a full simulation was made of the mSUGRA model with the corresponding set of parameters.

Process	Cross Section (pb)
Di-jet + μ + 0 partons	$8.4 \cdot 10^5$
Di-jet + μ + 1 partons	$8.2 \cdot 10^5$
Di-jet + μ + 2 partons	$2.2 \cdot 10^5$
Di-jet + μ + 3 partons	$2.9 \cdot 10^4$
Di-jet + μ + 4 partons	$2.0 \cdot 10^3$

Table 7.4: The LO cross section for the QCD+μ PYTHIA di-jet samples.

The new background samples were generated at the collision energy of 7 TeV. The CTEQ66 [11] set was used for the proton parton density function. A new detector description model was used with updated detector geometry based on the early 2010 collision data. The GEANT4 [52] program was used to simulate the detector response.

When proton beams are collided at high intensity, as the LHC is designed to do, the probability of more than one pair of protons colliding during one bunch crossing becomes non-negligible. When designing an analysis this effect (called pileup) must be taken into consideration, in order to be applicable to the high intensity environment at the LHC. The MC samples used in this analysis take pileup into account, by adding a number of minimum bias collisions to each simulated event. The extra vertices are reweighted to match the average number of vertices found in data.

7.2.4 Event selection

The definitions of the physics objects were previously described in section 3. They differ from the definitions used in the 10 TeV MC analysis, incorporating improved knowledge of the detector, and dealing with experimental challenges of analyzing actual data.

To get a good signal over background ratio, a set of cuts was applied to the events passing the trigger. A summary of these cuts is listed in Table 7.5. Events were selected with exactly one muon with $p_T > 20\,\mathrm{GeV}$, reducing the QCD background considerably. Events with a second lepton (electron or muon) with $p_T > 10\,\mathrm{GeV}$ were discarded, reducing the Z+jets and dileptonic $t\bar{t}$ background. This second lepton veto also eliminates the overlap with other SUSY analyses in the ATLAS collaboration requiring two or more leptons. Four jets with $|\eta| < 2.5$ and $p_T > 30\,\mathrm{GeV}$ were required, with a requirement for the leading jet to have $p_T > 60\,\mathrm{GeV}$. Requiring four jets means the mass of the hadronicaly decaying top can be reconstructed, thus separating these events from other backgrounds and signal. An important feature of the method described in this chapter is that it not only distinguishes background from signal, but also the different background contributions, giving a valuable cross check on the validity of the result. The cut on the $E_{\mathrm{T}}^{\mathrm{miss}}/M_{\mathrm{eff}}$ ratio greatly reduces the number of QCD events in the data sample.

The signal and control region are defined in the ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) plane. The full region in the ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) plane where the background model is to be applied is defined by $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 900\,\mathrm{GeV}$ and $40 < M_{\mathrm{T}} < 900\mathrm{GeV}$. This region (denoted by the full region throughout this chapter) is split in a control region (CR) and a signal region (SR). The upper edge of the control region is defined by $E_{\mathrm{T}}^{\mathrm{miss}} < 125\,\mathrm{GeV}$ or $M_{\mathrm{T}} < 100\,\mathrm{GeV}$, thus defining an L-shaped region, shown in Figure 7.1. The signal region is defined by $E_{\mathrm{T}}^{\mathrm{miss}} > 125\,\mathrm{GeV}$ and $M_{\mathrm{T}} > 100\,\mathrm{GeV}$. Finally events in the signal region were required to have $M_{\mathrm{eff}} > 500\,\mathrm{GeV}$.

Description	Selection
Nr. of muons $p_T > 20\,\text{GeV}$	$N_\mu^{20} = 1$
Nr. of muons $p_T > 10\,\text{GeV}$	$N_\mu^{10} = 1$
Nr. of electrons $p_T > 10\,\text{GeV}$	$N_e^{10} = 0$
Nr. of Jets $p_T > 60\,\text{GeV}$	$N_j^{60} \geq 1$
Nr. of Jets $p_T > 30\,\text{GeV}$	$N_j^{30} \geq 4$
$E_\mathrm{T}^\mathrm{miss}$	$E_\mathrm{T}^\mathrm{miss} > 30\,\text{GeV}$
M_T	$M_\mathrm{T} > 40\,\text{GeV}$
M_eff	$E_\mathrm{T}^\mathrm{miss}/M_\mathrm{eff} > 0.25$
Control region	$M_\mathrm{T} < 100\,\text{GeV} \vee E_\mathrm{T}^\mathrm{miss} < 125\,\text{GeV}$
Signal region	$M_\mathrm{T} > 100\,\text{GeV} \wedge E_\mathrm{T}^\mathrm{miss} > 125\,\text{GeV}$, $M_\mathrm{eff} > 500\,\text{GeV}$

Table 7.5: Event selection criteria

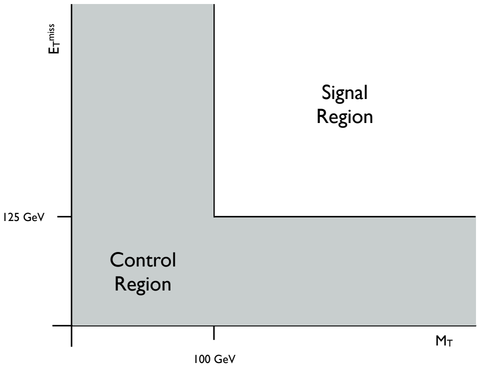


Figure 7.1: Definition of the control and signal regions in the $M_\mathrm{T}/E_\mathrm{T}^\mathrm{miss}$ plane

The cuts on M_T and $E_\mathrm{T}^\mathrm{miss}$ were designed to reduce the W+jets and $t\bar{t}$ background. The M_eff cut is designed to enhance the signal over background ratio for the gluino and squark masses that are expected to be probed by this method. These cuts were adopted from the ATLAS SUSY analysis [67].

The cut on M_eff poses an issue for the analysis presented here. Applying the same cut on the CR would reduce the background statistics to such a degree that it would be impossible to find the shape of the distributions. Thus an extra dimension has been added to the problem of extrapolation, requiring a fourth dimension to be added to the model. This dimension would add extra shape parameters to the model, and be strongly correlated to the $E_\mathrm{T}^\mathrm{miss}$ dimension. To keep the model simple, the efficiency of the M_eff cut is not a parameter of the fit, but is taken directly from MC.

A study of the expected significance attainable by this method averaged over all SUSY points for different values of the $E_\mathrm{T}^\mathrm{miss}$ and M_T boundary of the signal region shows that the effect of the M_eff cut could possibly be reproduced by posing a more stringent cut on $E_\mathrm{T}^\mathrm{miss}$, namely $E_\mathrm{T}^\mathrm{miss} \gtrsim 200\,\text{GeV}$. This would restore the data driven policy of the method. The cuts of the ATLAS SUSY group were adopted here in order to retain some measure of comparability

Component	Event count	Corrected count
Full Range		
$T1$	25.3	
W	11 .2	
$T2$	2.8	
TP	7.2	5.4
CW	29 .2	30.5
total MC	40	
data	46	
Control region		
$T1$	25.0	
W	10.9	
$T2$	2.5	
TP	7.1	5.3
CW	28.7	30.5
total MC	39	
data	43	
Signal region		
$T1$	0.1	
W	0.2	
$T2$	0.3	
TP	0.06	0.05
CW	0.3	0.3
total MC	0.6	
data	1	

Table 7.6: The event counts for both data and MC simulated samples of the SM backgrounds in the full, control and signal region. The MC counts are normalized to an integrated luminosity of $35\,\mathrm{pb}^{-1}$. The $t\bar{t}$ sample is split in $T1$ and $T2$. The $T1$ sample is split in TP and TC. The TC sample is added to the W sample to form CW. The yields of the CW (TP) sample in the third column is used to compare with the fitted yields. These are corrected for contamination by TP (TC) events.

to other results of the group.

Table 7.6 shows the expected number of events for the background MC samples normalized to an integrated luminosity of $35\,\mathrm{pb}^{-1}$. The same table shows the measured number of events in data. The numbers are quoted for the full ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) range, the signal region and the control region.

7.3 The combined fit model for first data

Since the dataset used for this analysis contains much fewer events than the one used in the previous MC based analysis, one expects that a somewhat simplified model can be used to describe the data. It turns out that the correlations between observables are small with respect

Name	Description	Comment
TT	Leptonic $t\bar{t}$	generated using MC@NLO/ALPGEN
$T1$	Semileptonic $t\bar{t}$	Split from TT on generator level
$T2$	Dileptonic $t\bar{t}$	Split from TT on generator level
TP	Top peak	Split from $T1$ using method of section 6.4.3.
TC	Top combinatorics	Split from $T1$ using method of section 6.4.3.
W	W+jets	Generated using ALPGEN
CW	W plus TC	

Table 7.7: The different definitions of the background components.

	$E_{\mathrm{T}}^{\mathrm{miss}}$	M_{T}	M_{jjj}
TP	$\mathrm{TTComb}(E_{\mathrm{T}}^{\mathrm{miss}};\mu,\sigma,\alpha)$	$\mathrm{TTComb}(M_{\mathrm{T}};\mu,\sigma,\alpha)$	$\mathrm{Gaussian}(M_{\mathrm{jjj}};\mu,\sigma)$
CW	$\mathrm{TTComb}(E_{\mathrm{T}}^{\mathrm{miss}};\mu,\sigma,\alpha)$	$\mathrm{TTComb}(M_{\mathrm{T}};\mu,\sigma,\alpha)$	$\mathrm{TTComb}(M_{\mathrm{jjj}};\mu,\sigma,\alpha)$
$T2$	$\mathrm{TTComb}(E_{\mathrm{T}}^{\mathrm{miss}};\mu,\sigma,\alpha)$	$\mathrm{TTComb}(M_{\mathrm{T}};\mu,\sigma,\alpha)$	$\mathrm{TTComb}(M_{\mathrm{jjj}};\mu,\sigma,\alpha)$

Table 7.8: The one-dimensional components of the background model. In principle parameter is unique to each component. Exceptions to this will be discussed in the text. The CW and TP M_{jjj} models are listed here as used in the combined fit. To determine the shape parameters in the prefit, the contamination of TC events in the TP sample and TP events in the CW sample is taken into account.

to the statistical uncertainty on the parameters, and can be ignored. This section will describe the new model, including a new SUSY Ansatz. Again, the $T1$ sample is split in a peak and a combinatorics part. The top peak (TP) part is modeled separately, and the combinatorics (TC) is added to the W sample to form a CW sample. Table 7.7 summarizes the definition and naming of the different samples.

7.3.1 Background model

Figure 7.2 shows the non-conditional model for all three backgrounds: W+combinatorics (CW), $t\bar{t}$ peak (TP) and dileptonic $t\bar{t}$ ($T2$), compared to the MC simulated data for each respective background. These data samples passed through the event selection of the previous section, where data in the full ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) region (control and signal region combined) was considered. This means the background model is made to describe the background in both control and signal region, a necessary feature if it is to be used to extrapolate between these two regions.

The model is shown projected on the $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj} axes. Note that the $T2$ background has very low statistics, caused by the second lepton veto of this analysis: only dileptonic $t\bar{t}$ events where one of the lepton was not identified pass this cut.

Table 7.8 shows the one-dimensional components from which the total combined fit model was constructed. With the exception of the TP M_{jjj} model, all models are described by the TTComb shape defined in the previous chapter. The TTComb is an exponential function of

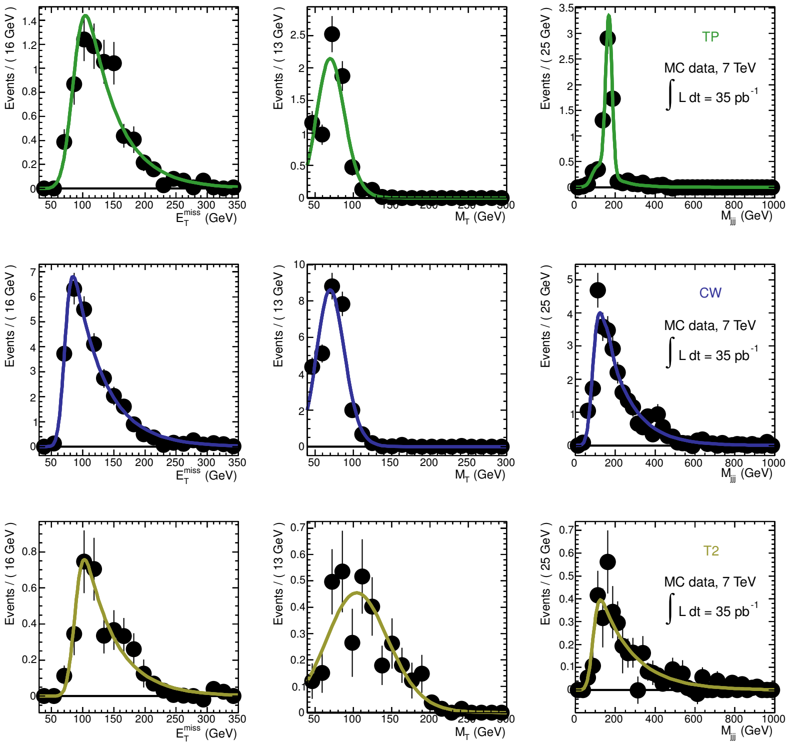


Figure 7.2: The model for all three backgrounds, projected on the three observables' axes. Black dots show the simulated events for each background sample, the solid lines are the non-conditional model.

the observable x at high values of x, with a smooth cutoff to zero at low values of x. [1]

Each three-dimensional component model is created from the one-dimensional components by multiplication:

$$F_{BG}^{3D}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) = F_{BG}^{1D}(E_{\mathrm{T}}^{\mathrm{miss}}) \times F_{BG}^{1D}(M_{\mathrm{T}}) \times F_{BG}^{1D}(M_{\mathrm{jjj}}) \tag{7.2}$$

The total background model now looks like:

$$\begin{aligned} F_{3D}^{comb}(E_{\mathrm{T}}^{\mathrm{miss}}, M_{\mathrm{T}}, M_{\mathrm{jjj}}) &= (1 - f_{CW})\left[f_{T2} \cdot F_{T2}^{3D} + (1 - f_{T2}) \cdot F_{TP}^{3D}\right] + \\ &\quad f_{CW} \cdot F_{CW}^{3D} \end{aligned} \tag{7.3}$$

with f_{CW} the fitted fraction of CW in the total background, and f_{T2} the fitted fraction of dileptonic $t\bar{t}$ in the total $t\bar{t}$ contribution.

---

[1] The functional form of the TTComb is:

$$\mathrm{TTComb}(x; \mu, \sigma, \alpha) = \frac{1}{N}(1 + erf(x; \mu, \sigma)) \times e^{-\alpha x} \tag{7.1}$$

There are a few noteworthy features when comparing this model to the one used in the previous chapter. The first feature is a simplification of the M_{T} model. The M_{T} distribution of TP, $T2$ and CW is described by a single TTComb, where previously an exponential tail and two Gaussians were used. The TP M_{jjj} model is a Gaussian, which describes the top mass peak.

To get a first estimate of the shape parameters of the combined fit model, each 3D component is fitted to the corresponding background sample. Thus the $T2$ model is fitted to a sample with only $T2$ events, and the CW model is fitted to a sample of W+jets and top combinatorics events. In the prefit stage , the TP M_{jjj} model is expanded with a TTComb to account for contamination of TC events in the TP sample. Likewise, a Gaussian is added to the CW model to account for TP contamination. After the prefit the combined model with the estimated shape parameter values can be fitted to a sample containing all backgrounds. This can be either a combined simulated sample of all backgrounds and a SUSY model, or to a dataset of recorded events. In this stage, the contamination fraction of both the TP and the CW components is set to 0, such that the TP model is sensitive only to the Gaussian mass peak in the M_{jjj} distribution of the dataset. All top events where the hadronic top mass is not reconstructed are accounted for by the CW model.

The naming of the model parameters is denoted as < `source`>_< `observable`>_< `parameter`>. Thus the μ of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb of the CW model is named `CW_etmiss_mean`. The shape parameter of the TTComb exponential tail is named `base`. The peak fraction in the TP and CW model is named `TP(CW)_mjjj_fpeak`. This naming will be used throughout this chapter. Finally all parameters whose name starts with `G` (for "global") are shared between to models, as will be explained in section 7.3.2.

7.3.2 Global shape parameters

In principle the parameter of each 1D component model of Table 7.8 is unique to that component. Hence the α of the TP M_{T} model is different from the α of the CW M_{T} or the TP $E_{\mathrm{T}}^{\mathrm{miss}}$ model. From a physics viewpoint there are exceptions to this rule that can be made. One example is the TP and CW M_{T} distribution, which is largely determined in both cases by the mass of the W that decays to a lepton and a neutrino.

To accommodate this behavior the model was setup such that shape parameters can be shared between 1D components. Corresponding parameters of the different components (e.g. the α parameter of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb, or the μ parameter of the TP and CW M_{T} distribution) were compared. When these parameters of different models were fitted in the prefit stage with values within 0.5σ of each other they were designated shared parameters. These shared, or global parameters are shown in Table 7.9.

7.3.3 Conditional PDFs

In the previous chapter correlations between M_{T}, M_{jjj} and $E_{\mathrm{T}}^{\mathrm{miss}}$ needed to be taken into consideration using conditional product PDFs. The correlations between the observables could be seen when binning the data in one dimension, and examining the shape of the distributions in the other two. In this section, using the CW M_{T} model it will be shown that these correlations are now small compared to the statistical errors, and can be ignored. The plots used in this section are also made for the other models, and shown in appendix B.

parameter	description	shared between
`G_mtrans_mean`	mean of M_{T} TTComb	CW, TP, $T2$
`G_mtrans_sigma`	width of M_{T} TTComb	CW, TP
`G_mtrans_base`	slope of M_{T} TTComb	CW, TP
`G_etmiss_base`	slope of $E_{\mathrm{T}}^{\mathrm{miss}}$ TTComb	CW, TP
`TP_mjjj_mean`	mean of the top mass peak	CW, TP
`TP_mjjj_sigma`	width of the top mass peak	CW, TP
`CW_mjjj_comb_mean`	mean of the CW M_{jjj} TTComb	CW, TP
`CW_mjjj_comb_sigma`	width of the CW M_{jjj} TTComb	CW, TP

Table 7.9: The shape parameters of the model that are shared between component models. The mean and width of the TP mass peak is used for the prefit stage only in the CW model to determine the TP contamination. Similarly, the mean and width of the CW M_{jjj} TTComb is used to determine TC contamination in TP.

The CW M_{T} model is a TTComb:

$$CW_{M_{\mathrm{T}}} = \mathrm{TTComb}(M_{\mathrm{T}}, \mu_{M_{\mathrm{T}}}^{CW}, \sigma_{M_{\mathrm{T}}}^{CW}, \alpha_{M_{\mathrm{T}}}^{CW}) \tag{7.4}$$

Figure 7.3 shows the M_{T} distribution of the W+jets MC sample in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$ as black dots. The solid line is the CW model fitted to the data in that bin. As a reference, the same model was fitted to all data, unbinned in $E_{\mathrm{T}}^{\mathrm{miss}}$, and projected on each bin. This model is shown in the same figure as the dashed line. It is clear that the shape of the M_{T} distribution does not change significantly as a function of $E_{\mathrm{T}}^{\mathrm{miss}}$. To illustrate this Figure 7.4 shows the fitted values of the floating parameters ($\mu_{M_{\mathrm{T}}}^{CW}$, $\sigma_{M_{\mathrm{T}}}^{CW}$ and $\alpha_{M_{\mathrm{T}}}^{CW}$) of the model in each bin of $E_{\mathrm{T}}^{\mathrm{miss}}$. The values of the parameter are scattered within $\sim 1\ \sigma$ around a central value. Note that for the sixth plot, where the dashed and solid line are quite different, the value of the width of the TTComb has a very large error. Similar analysis of the other 1D models are shown in appendix B. None of these models show significant correlations between either M_{T} or M_{jjj} and $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus the model used for this analysis of the first ATLAS data does not take correlations between observables into account. That leaves us with a model for the background with 17 independent shape parameters, and two recursive fraction parameters.

7.3.4 SUSY Ansatz

The last part that needs to be added to the model is an Ansatz shape for the SUSY contamination in the control region. The upper bound for the low $E_{\mathrm{T}}^{\mathrm{miss}}$ leg of the control region is set so low, that there are very few SUSY events expected in that region for 35 pb^{-1}, irrespective of the mSUGRA point under investigation. Therefore the Ansatz will be shown only in the region $E_{\mathrm{T}}^{\mathrm{miss}} > 40\,\mathrm{GeV}$ and $30 < M_{\mathrm{T}} < 100\,\mathrm{GeV}$. To illustrate the Ansatz, and the results of this model in the next sections, three points have been picked out of the grid in mSUGRA phase space defined in section 7.2.3:

- Point 1: $m_0 = 160\,\mathrm{GeV}$, $m_{1/2} = 175\,\mathrm{GeV}$, ratio = 1.3, N = 19, $\sigma_{NLO} = 30\,\mathrm{pb}$
- Point 2: $m_0 = 200\,\mathrm{GeV}$, $m_{1/2} = 145\,\mathrm{GeV}$, ratio = 0.5, N = 35, $\sigma_{NLO} = 64\,\mathrm{pb}$
- Point 3: $m_0 = 120\,\mathrm{GeV}$, $m_{1/2} = 130\,\mathrm{GeV}$, ratio = 0.6, N = 51, $\sigma_{NLO} = 140\,\mathrm{pb}$

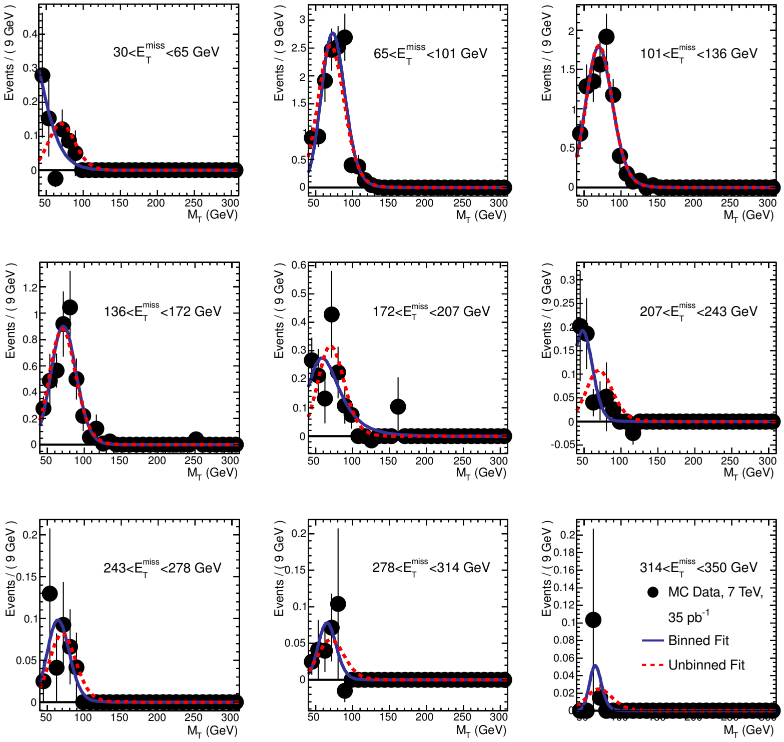


Figure 7.3: The M_T distribution of the W+jets background sample in bins of E_T^{miss}. The solid line is the non-conditional W M_T model fitted to the data in that bin. The dashed line is the model fitted to all data. The plot shows this model projected on the data in each E_T^{miss} bin.

where the ratio is defined as the number of events in the signal region divided by the number of events in the control region (N_{SIG}/N_{CR}), and N is the total number of signal events after selection in the full (control + signal) region . These points have different ratios and cross sections, and can serve well to illustrate the model with different SUSY signatures. Point 1 has the highest mass scale, with both the squark and gluino masses around 400 GeV. Point 3 has a much lower mass scale, and a very high cross section, with gluino and squark masses around 300 GeV. Not surprisingly it will become clear that the lower the SUSY mass scale, the more difficult it becomes to distinguish it from SM background.

The Ansatz used is quite simple: a Gaussian for the E_T^{miss} distribution, a TTComb for the M_{jjj} distribution, and an exponential function for the M_T distribution. Figure 7.5 shows this Ansatz for the three showcase points. Table 7.10 shows the parameters of this model.

This model with 6 shape parameters was fitted to all points of the SUSY grid. In the previous chapter it was shown that the value of the shape parameters of the SUSY Ansatz can be correlated. This correlation can be used to reduce the number of parameters in the Ansatz model.

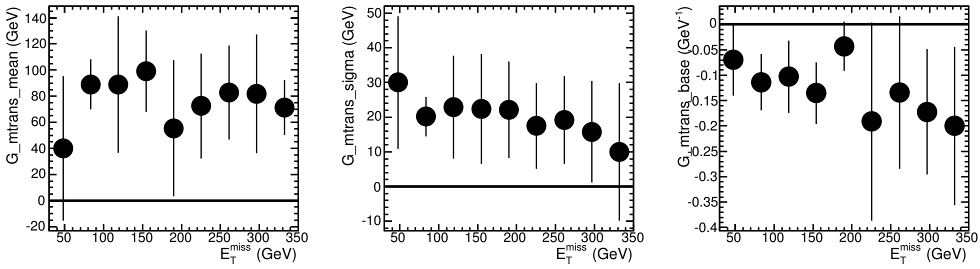


Figure 7.4: The fit parameters of the fit of the non-conditional W M_T model in bins of $E_\mathrm{T}^\mathrm{miss}$.

name	description
`SU_etmiss_mean`	Mean of $E_\mathrm{T}^\mathrm{miss}$ Gaussian
`SU_etmiss_sigma`	Width of $E_\mathrm{T}^\mathrm{miss}$ Gaussian
`SU_mtrans_mean`	Mean of M_T TTComb
`SU_mtrans_sigma`	Width of M_T TTComb
`SU_mtrans_base`	Decay rate of M_T tail
`SU_mjjj_base`	Decay rate of M_jjj exponential

Table 7.10: The six parameters of the SUSY Ansatz model. The five parameters starting from row number two will be expressed in terms of the first parameter, `SU_etmiss_mean`.

Figure 7.6 shows the shape parameters (averaged over all SUSY points) as a function of the mean of the $E_\mathrm{T}^\mathrm{miss}$ Gaussian (`SU_etmiss_mean`). A clear correlation between the parameters can be seen. The solid line shows the expression which replaced the parameter in the final Ansatz model. A linear expression was used for all parameters, where the expression for the base of the M_T exponential was limited to be lower than -0.001. The expression for `SU_mjjj_sigma` was fitted to the first four points only. The last three points are shifted downwards by a few outliers with low associated errors.

Figure 7.5 shows the Ansatz model with all shape parameters floating (solid line), and with five parameters replaced by a function of `SU_etmiss_mean` (dashed line). The three distinct showcase points can all be described by either model. Thus adding the SUSY Ansatz to the combined fit model adds two floating parameters: one SUSY yield and one SUSY shape parameter.

7.4 Test on simulated data sample

To gain insight on the performance of this method on the available dataset, first a series of test fits will be performed on a simulated data sample, which has the same size as the ATLAS 2010 dataset in terms of integrated luminosity. First a prefit will be performed to get a first estimate of the shape parameters. Each 3D background component model will be fitted to simulated data for the corresponding background. The 1D projections of the three background components after the prefit were already shown in Figure 7.2. The shape parameters of the three components with their fitted value are shown in Table 7.11.

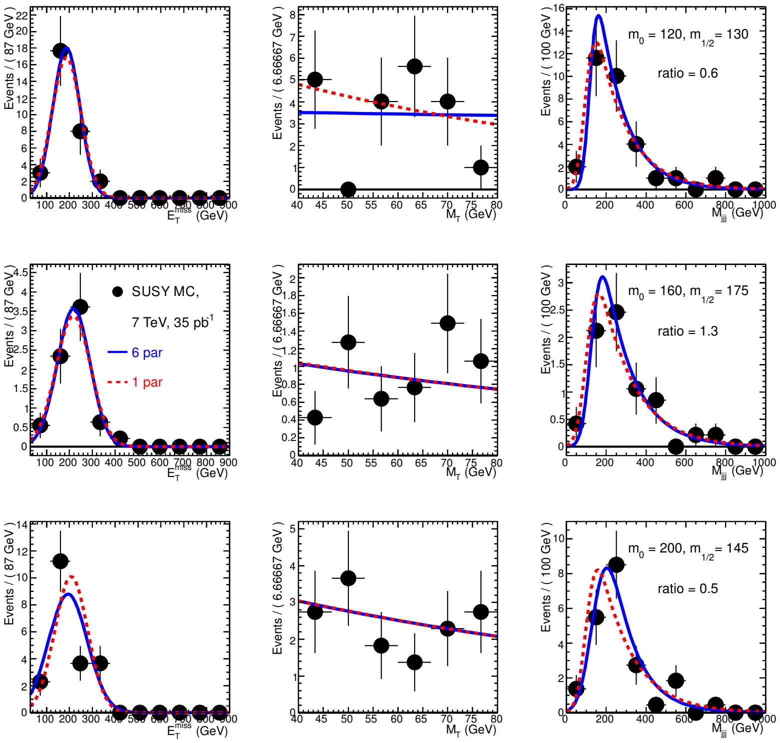


Figure 7.5: The SUSY Ansatz for the three showcase points. The solid line shows the model with all parameters floating, the dashed line shows the model with 5 parameters replaced by functions of `SU_etmiss_mean`. The numbers for m_0 and $m_{1/2}$ in the plots are in units of GeV.

The shape parameters of the model will then be fixed to the fitted value, with the exception of the TP and CW M_{jjj} peak fractions, which will be fixed at 1 and 0 respectively. Thus the TP model only looks for the hadronic top mass peak in the M_{jjj} distribution , and all top combinatorics will be accounted for by the CW model. The samples that were labeled TP and CW of course still contain the contamination of TC and TP events. To correct for this contamination the contamination fraction that was used for this correction was taken directly from the prefit. All table in the following sections display the corrected MC expectation for TP and CW. Table 7.6 shows the number of simulated events for each background component, rescaled to $35\,\text{pb}^{-1}$ integrated luminosity.

The first fit will be performed on a background-only simulated dataset in the full ($E_{\text{T}}^{\text{miss}}$, M_{T}) region, comparing the fitted yields to the input yields of the dataset.

The second test will be fitting the model in the CR, and extrapolating to the SR. The extrapolated yields for each background will be multiplied by the efficiency of the $M_{\text{eff}} > 500\,\text{GeV}$ cut taken from MC. This efficiency is 0.4 for the TP model, 0.5 for CW and 0.3 for $T2$.

For the last test of the background-only model a number of shape parameters will be

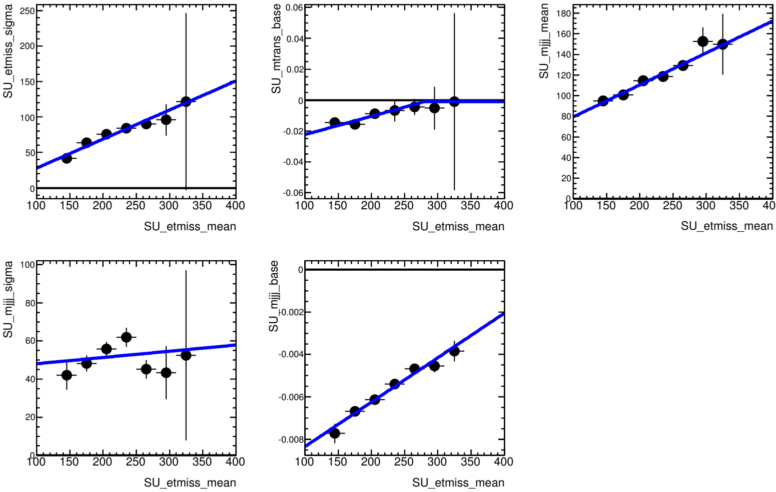


Figure 7.6: Profile plot of five correlated SUSY Ansatz parameters as a function of `SU_etmiss_mean`. The line is the fitted expression that was substituted for that parameter in the SUSY Ansatz model.

released in the combined fit, based on the systematic uncertainty induced by fixing these parameters.

When the maximum set of floating shape parameters is determined, the SUSY Ansatz will be added to the model. This complete combined fit model will be fitted to a dataset without SUSY, and to three datasets containing simulated signal from one of the three showcase points.

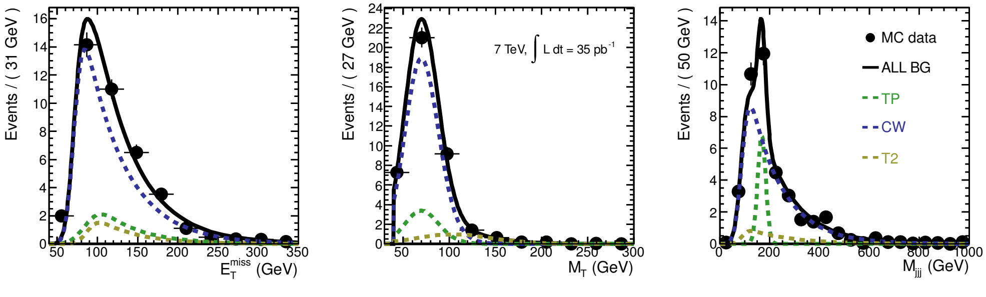


Figure 7.7: Background model fitted to a simulated background-only sample in the full ($E_{\mathrm{T}}^{\mathrm{miss}}$,$M_{\mathrm{T}}$) region.

Parameter	Fitted value	
`CW_etmiss_mean`	73 ± 7.6	GeV
`CW_etmiss_sigma`	12 ± 7.6	GeV
`CW_mjjj_base`	$(-7.6 \pm 2.3) \cdot 10^{-3}$	GeV^{-1}
`CW_mjjj_peak_sigma`	35 ± 31	GeV
`G_etmiss_base`	$(-20 \pm 3.8) \cdot 10^{-3}$	GeV^{-1}
`G_etmiss_mean`	90 ± 13	GeV
`G_mjjj_mean`	91 ± 19	GeV
`G_mjjj_sigma`	31 ± 15	GeV
`G_mtrans_base`	$(-11 \pm 4.3) \cdot 10^{-2}$	GeV^{-1}
`G_mtrans_mean`	90 ± 18	GeV
`G_mtrans_sigma`	23 ± 4.7	GeV
`T2_etmiss_sigma`	14 ± 22	GeV
`T2_mjjj_base`	$(-6.0 \pm 3.5) \cdot 10^{-3}$	GeV^{-1}
`T2_mtrans_mean`	$1.1 \cdot 10^{2} \pm 26$	GeV
`T2_mtrans_sigma`	41 ± 20	GeV
`TP_etmiss_sigma`	20 ± 20	GeV
`TP_mjjj_base`	$(-1.2 \pm 1.6) \cdot 10^{-2}$	GeV^{-1}
`TP_mjjj_fpeak`	$(6.9 \pm 3.6) \cdot 10^{-1}$	
`TP_mjjj_peak_mean`	$1.7 \cdot 10^{2} \pm 11$	GeV
`TP_mjjj_peak_sigma`	16 ± 9.6	GeV

Table 7.11: All parameters of the fit, plus the value fitted in the prefit.

7.4.1 Background only sample

Fitting in full range

Figure 7.7 shows the background model (black solid line) fitted to a simulated data sample (black dots). The colored dashed lines are the contributions from the different backgrounds. Only the background yield parameters and the top M_{jjj} mass parameters were fitted, and the fit was performed in the full ($E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T}) region. The resulting yields are shown in Table 7.12. As expected, the yields are fitted very well. The figure next to the table displays the full correlation matrix, which shows that correlations between the fitted parameters are well under control.

Fitting in control region and extrapolation

The next step is to fit the background model in the control region and extrapolate it to the signal region. The result is shown in Table 7.13. As most of the background events are in the control region, there should not be much difference between this fit and the previous one. Again, the correlation matrix is shown next to the table.

7.4.2 Yield + shape parameters

For the final test of the background model some shape parameters of the model were allowed to fluctuate in the fit. The choice of which parameters were to be released in the fit was based

Sample	Fitted yield	True yield
TP	5.5 ± 3.8	5.4
CW	31 ± 4.7	30.5
T2	3.6 ± 2.8	3.4

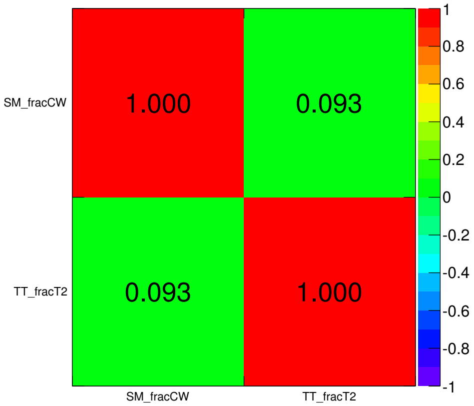


Table 7.12: Fitted yields of the background-only model to a background-only sample in the full ($E_{\mathrm{T}}^{\mathrm{miss}}$,$M_{\mathrm{T}}$) region. The figure shows the correlation between the parameters `SM_fracCW` and `TT_fracT2`, the recursive fractions of equation 7.3.

Sample	Fitted yield	True yield
	Control region	
TP	5.4 ± 3.8	5.3
CW	30 ± 4.8	30.5
T2	2.7 ± 2.9	2.5
	Signal region	
TP	0.08 ± 0.05	0.05
CW	0.4 ± 0.06	0.3
T2	0.3 ± 0.4	0.3

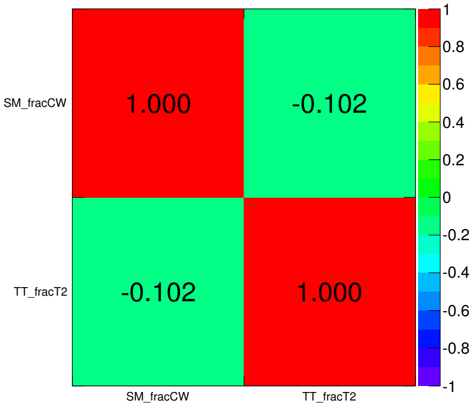


Table 7.13: Fitted yields of the background-only model to a background-only sample in the control region, and the extrapolated yields in the signal region.

on the systematic error induced by fixing this parameter. Here the result of Table 7.23 is anticipated. This table uses the change in the expected number of events by shifting shape parameters to estimate the systematic uncertainty induced by fixing that parameter.

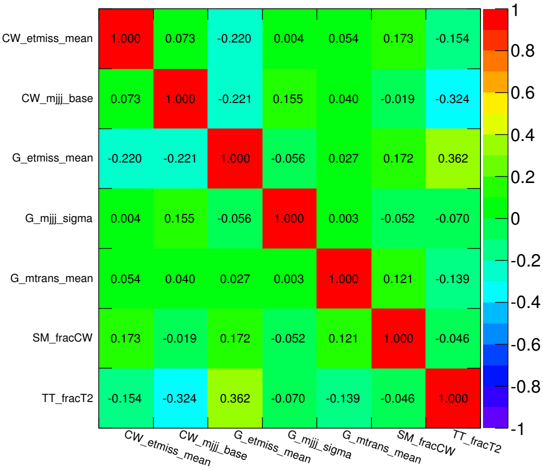
Based on the results on the systematic errors, the parameters that induced the largest systematic errors were allowed to float in the fit one by one. After a parameter was released, correlations with other parameters were checked. If these correlations were large, the parameter was fixed at the prefit value, as these correlations indicate a redundancy in the fit. This redundancy should lead to a smaller systematic error associated to the fixed parameter when the correlated parameter is allowed to float.

Floating parameter
`CW_etmiss_mean`
`CW_mjjj_base`
`G_etmiss_mean`
`G_mjjj_sigma`
`G_mtrans_mean`
Fixed due to correlations
`CW_etmiss_sigma`
`G_mjjj_mean`
`G_mtrans_base`
`G_mtrans_sigma`

Table 7.14: The top half of this table shows the shape parameters of the background model that are allowed to fluctuate in the fit. The bottom half shows the parameters that are fixed due to large correlations with the other parameters.

The five parameters in the top half of Table 7.14 were allowed to float in the fit. The four parameters in the bottom half were fixed due to correlations, but had a significantly reduced associated systematic error. Note that all $T2$ shape parameters were fixed. Any floating $T2$ shape parameter confused the fitting procedure when some SUSY signal was added to the simulated sample (see the next section). Floating more parameters induced instabilities in the fit, specifically trying to assign values to shape parameters that lie outside a reasonable range,

Sample	Fitted yield	True yield
Control region		
TP	5.0 ± 4.0	5.3
CW	31 ± 5	30.5
T2	2.7 ± 3.2	2.5
Signal region		
TP	0.08 ± 0.06	0.05
CW	0.4 ± 0.2	0.3
T2	0.4 ± 0.5	0.3
Fitted shape parameters		
`CW_etmiss_mean`		73 ± 4
`CW_mjjj_base`		$(-8 \pm 1) \cdot 10^{-3}$
`G_etmiss_mean`		93 ± 16
`G_mjjj_sigma`		32 ± 10
`G_mtrans_mean`		91 ± 4



	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	TT_fracT2
CW_etmiss_mean	1.000	0.073	-0.220	0.004	0.054	0.173	-0.154
CW_mjjj_base	0.073	1.000	-0.221	0.155	0.040	-0.019	-0.324
G_etmiss_mean	-0.220	-0.221	1.000	-0.056	0.027	0.172	0.362
G_mjjj_sigma	0.004	0.155	-0.056	1.000	0.003	-0.052	-0.070
G_mtrans_mean	0.054	0.040	0.027	0.003	1.000	0.121	-0.139
SM_fracCW	0.173	-0.019	0.172	-0.052	0.121	1.000	-0.046
TT_fracT2	-0.154	-0.324	0.362	-0.070	-0.139	-0.046	1.000

Table 7.15: Fitted yields in control and signal region for the background model with floating shape parameters. The `mean` and `sigma` parameters are in units of GeV, the `base` parameters in units of GeV^{-1}.

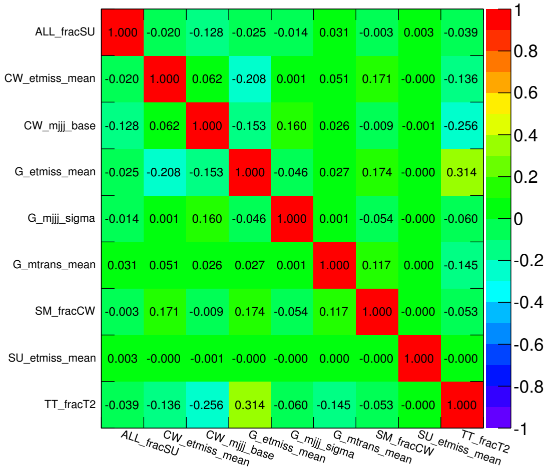
even when fitting to a simulated dataset (where the correct shape of the model is known).

The result of this fit is shown in Table 7.15. No significant changes are found in the fitted yield parameters with respect to the fit with fixed shapes. As expected, the uncertainties on the yield estimates are somewhat larger.

7.4.3 Background + SUSY sample

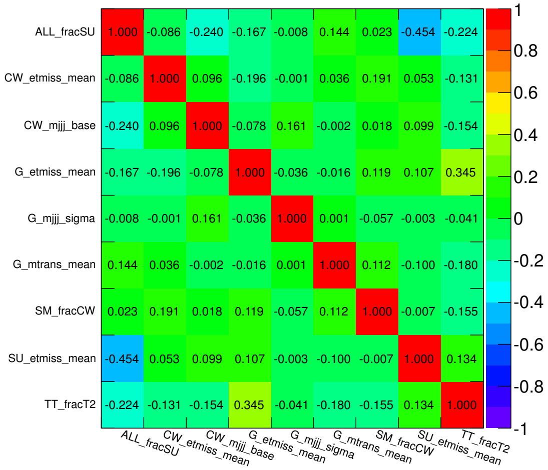
To test the SUSY Ansatz, the complete model (background + SUSY Ansatz) was fitted to a simulated dataset with no SUSY component, and with the data from the three showcase points added. The result is shown in Tables 7.16 and 7.17. The first table, with no SUSY in the sample shows that the method is unbiased: no SUSY is found in a sample without SUSY. For the first showcase point the fitted yields correspond nicely to the input values. When data is added for the second and third showcase point however, the estimated yields of the SM backgrounds are off. This behavior remains when the shape parameters of the background model are fixed to the prefit value. This means that for regions in mSUGRA phase space that look like showcase points 2 or 3 the CW and $T2$ yields are over-estimated. Analysis of the other points in the grid showed that in general the exact number of SUSY events in the control region is harder to determine with decreasing ratio, or increasing SUSY cross section. This is something that has to be considered when performing the fit on data.

No SUSY in sample		
Sample	Fitted yield	True yield
TP	5.1 ± 4.0	5.3
CW	30 ± 5	30.5
T2	2.6 ± 3.0	2.5
SU	0.7 ± 0.9	0
Signal region		
TP	0.08 ± 0.07	0.05
CW	0.4 ± 0.2	0.3
T2	0.3 ± 0.4	0.3
Fitted shape parameters		
`CW_etmiss_mean`		73 ± 4
`CW_mjjj_base`		$(-9 \pm 1) \cdot 10^{-3}$
`G_etmiss_mean`		92 ± 15
`G_mjjj_sigma`		31 ± 10
`G_mtrans_mean`		91 ± 4
`SU_etmiss_mean`		400 ± 298



	ALL_fracSU	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	SU_etmiss_mean	TT_fracT2
ALL_fracSU	1.000	-0.020	-0.128	-0.025	-0.014	0.031	-0.003	0.003	-0.039
CW_etmiss_mean	-0.020	1.000	0.062	-0.208	0.001	0.051	0.171	-0.000	-0.136
CW_mjjj_base	-0.128	0.062	1.000	-0.153	0.160	0.026	-0.009	-0.001	-0.256
G_etmiss_mean	-0.025	-0.208	-0.153	1.000	-0.046	0.027	0.174	-0.000	0.314
G_mjjj_sigma	-0.014	0.001	0.160	-0.046	1.000	0.001	-0.054	-0.000	-0.060
G_mtrans_mean	0.031	0.051	0.026	0.027	0.001	1.000	0.117	0.000	-0.145
SM_fracCW	-0.003	0.171	-0.009	0.174	-0.054	0.117	1.000	-0.000	-0.053
SU_etmiss_mean	0.003	-0.000	-0.001	-0.000	-0.000	0.000	-0.000	1.000	-0.000
TT_fracT2	-0.039	-0.136	-0.256	0.314	-0.060	-0.145	-0.053	-0.000	1.000

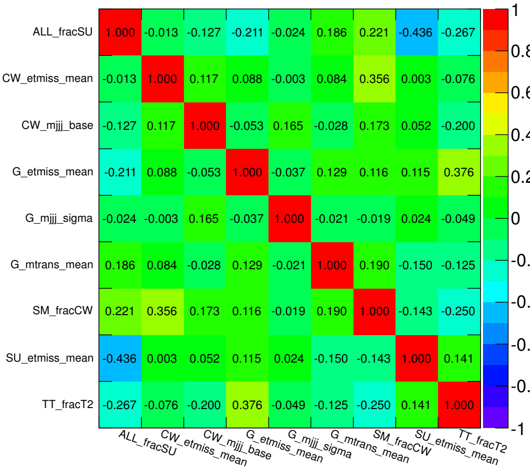
SUSY showcase point 1		
Sample	Fitted yield	True yield
TP	5.5 ± 4.1	5.3
CW	31 ± 6	30.5
T2	2.2 ± 3.1	2.5
SU	6.9 ± 4.2	7.6
Signal region		
TP	0.09 ± 0.07	0.05
CW	0.4 ± 0.2	0.3
T2	0.3 ± 0.5	0.3
Fitted shape parameters		
`CW_etmiss_mean`		74 ± 4
`CW_mjjj_base`		$(-9 \pm 2) \cdot 10^{-3}$
`G_etmiss_mean`		93 ± 16
`G_mjjj_sigma`		31 ± 10
`G_mtrans_mean`		91 ± 4
`SU_etmiss_mean`		238 ± 31



	ALL_fracSU	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	SU_etmiss_mean	TT_fracT2
ALL_fracSU	1.000	-0.086	-0.240	-0.167	-0.008	0.144	0.023	-0.454	-0.224
CW_etmiss_mean	-0.086	1.000	0.096	-0.196	-0.001	0.036	0.191	0.053	-0.131
CW_mjjj_base	-0.240	0.096	1.000	-0.078	0.161	-0.002	0.018	0.099	-0.154
G_etmiss_mean	-0.167	-0.196	-0.078	1.000	-0.036	-0.016	0.119	0.107	0.345
G_mjjj_sigma	-0.008	-0.001	0.161	-0.036	1.000	0.001	-0.057	-0.003	-0.041
G_mtrans_mean	0.144	0.036	-0.002	-0.016	0.001	1.000	0.112	-0.100	-0.180
SM_fracCW	0.023	0.191	0.018	0.119	-0.057	0.112	1.000	-0.007	-0.155
SU_etmiss_mean	-0.454	0.053	0.099	0.107	-0.003	-0.100	-0.007	1.000	0.134
TT_fracT2	-0.224	-0.131	-0.154	0.345	-0.041	-0.180	-0.155	0.134	1.000

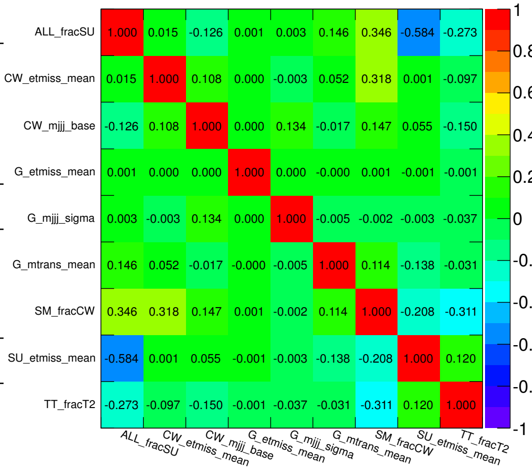
Table 7.16: Fitted background and signal yields for samples without, and with the first SUSY showcase point added. The parameter `ALL_fracSU` is the recursive fraction of SU events when the SUSY Ansatz is combined with the background model of equation 7.3. The `mean` and `sigma` parameters are in units of GeV, the `base` parameters in units of GeV^{-1}.

SUSY showcase point 2		
Sample	Fitted yield	True yield
TP	4.0 ± 3.9	5.3
CW	37 ± 7.3	30.5
T2	6.8 ± 6.6	2.5
SU	13 ± 5	23
Signal region		
TP	0.1 ± 0.1	0.05
CW	0.6 ± 0.3	0.3
T2	1.8 ± 2.0	0.3
Fitted shape parameters		
`CW_etmiss_mean`		75 ± 4
`CW_mjjj_base`		$(-9 \pm 2) \cdot 10^{-3}$
`G_etmiss_mean`		116 ± 33
`G_mjjj_sigma`		31 ± 9
`G_mtrans_mean`		93 ± 4
`SU_etmiss_mean`		238 ± 21



	ALL_fracSU	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	SU_etmiss_mean	TT_fracT2
ALL_fracSU	1.000	-0.013	-0.127	-0.211	-0.024	0.186	0.221	-0.436	-0.267
CW_etmiss_mean	-0.013	1.000	0.117	0.088	-0.003	0.084	0.356	0.003	-0.076
CW_mjjj_base	-0.127	0.117	1.000	-0.053	0.165	-0.028	0.173	0.052	-0.200
G_etmiss_mean	-0.211	0.088	-0.053	1.000	-0.037	0.129	0.116	0.115	0.376
G_mjjj_sigma	-0.024	-0.003	0.165	-0.037	1.000	-0.021	-0.019	0.024	-0.049
G_mtrans_mean	0.186	0.084	-0.028	0.129	-0.021	1.000	0.190	-0.150	-0.125
SM_fracCW	0.221	0.356	0.173	0.116	-0.019	0.190	1.000	-0.143	-0.250
SU_etmiss_mean	-0.436	0.003	0.052	0.115	0.024	-0.150	-0.143	1.000	0.141
TT_fracT2	-0.267	-0.076	-0.200	0.376	-0.049	-0.125	-0.250	0.141	1.000

SUSY showcase point 3		
Sample	Fitted yield	True yield
Control region		
TP	7.0 ± 4.2	5.3
CW	39 ± 7	30.5
T2	12 ± 8	2.5
SU	13 ± 7	33
Signal region		
TP	0.2 ± 0.1	0.05
CW	0.5 ± 0.2	0.3
T2	3.8 ± 2.5	0.3
Fitted shape parameters		
`CW_etmiss_mean`		75 ± 4
`CW_mjjj_base`		$(-10 \pm 2) \cdot 10^{-3}$
`G_etmiss_mean`		120 ± 5
`G_mjjj_sigma`		27 ± 8
`G_mtrans_mean`		90 ± 4
`SU_etmiss_mean`		230 ± 29



	ALL_fracSU	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	SU_etmiss_mean	TT_fracT2
ALL_fracSU	1.000	0.015	-0.126	0.001	0.003	0.146	0.346	-0.584	-0.273
CW_etmiss_mean	0.015	1.000	0.108	0.000	-0.003	0.052	0.318	0.001	-0.097
CW_mjjj_base	-0.126	0.108	1.000	0.000	0.134	-0.017	0.147	0.055	-0.150
G_etmiss_mean	0.001	0.000	0.000	1.000	0.000	-0.000	0.001	-0.001	-0.001
G_mjjj_sigma	0.003	-0.003	0.134	0.000	1.000	-0.005	-0.002	-0.003	-0.037
G_mtrans_mean	0.146	0.052	-0.017	-0.000	-0.005	1.000	0.114	-0.138	-0.031
SM_fracCW	0.346	0.318	0.147	0.001	-0.002	0.114	1.000	-0.208	-0.311
SU_etmiss_mean	-0.584	0.001	0.055	-0.001	-0.003	-0.138	-0.208	1.000	0.120
TT_fracT2	-0.273	-0.097	-0.150	-0.001	-0.037	-0.031	-0.311	0.120	1.000

Table 7.17: Fitted background and signal yields for samples with second and third SUSY showcase points added. The parameter `ALL_fracSU` is the recursive fraction of SU events when the SUSY Ansatz is combined with the background model of equation 7.3. The `mean` and `sigma` parameters are in units of GeV, the `base` parameters in units of GeV^{-1}.

7.5 Results on ATLAS 7 TeV data

	Control region	Signal region
data	43	1
MC	38	0.6

Table 7.18: Number of events in control and signal region as measured in data, compared to MC expectation.

Finally the combined fit method can be applied to the data measured by the ATLAS detector during 2010. This will give an estimate of the total background yield in the signal region, which will be used to either find evidence for an excess of events, or set a limit on the total number of new physics events in the signal region. This result will then be analyzed for the specific case of the mSUGRA models in the grid that was described earlier.

A first comparison of data and MC is given in Figure 7.8, which shows the distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj} of the the three MC background samples as a stacked histogram, with the measured distributions as black dots.

Table 7.18 shows the number of events in the control and signal region recorded in 2010, and compared to the total MC expectation from W+jets and $t\bar{t}$.

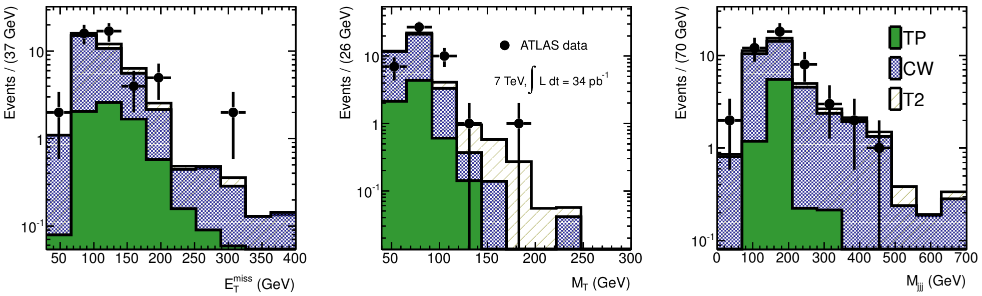
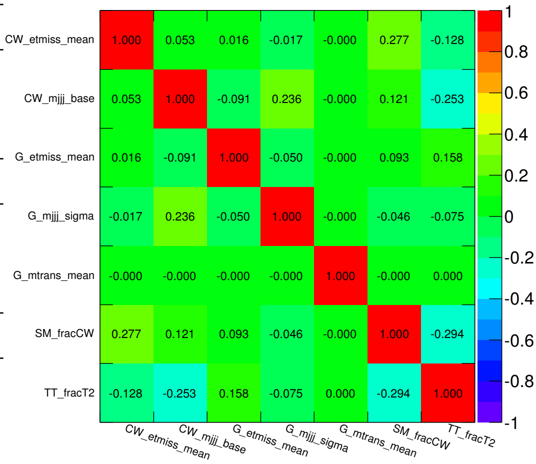


Figure 7.8: Comparison of the distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj} as measured in data to those as expected from MC.

7.5.1 Fitting the background model to data

Table 7.19 shows the fitted yields in the control region for the background-only model fitted to the data. All fitted yields agree with the MC expectation within statistical uncertainties. There are no inconsistencies between data and MC which would indicate either incorrect modeling of the SM in the simulation or contamination of non-SM signal in the control region. The same table shows the extrapolation of the model to the signal region, where the efficiency of the M_{eff} cut is again taken from MC, and applied in this table.

Sample	Fitted yield	MC expectation
Control region		
TP	7.5 ± 3.7	5.3
CW	30 ± 5	30.5
T2	5.2 ± 4.3	2.5
Signal region		
TP	0.3 ± 0.2	0.05
CW	0.8 ± 0.2	0.3
T2	1.0 ± 0.9	0.3
Fitted shape parameters		
`CW_etmiss_mean`	68 ± 4	
`CW_mjjj_base`	$(-9 \pm 2) \cdot 10^{-3}$	
`G_etmiss_mean`	106 ± 9	
`G_mjjj_sigma`	42 ± 11	
`G_mtrans_mean`	100 ± 4	



	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	TT_fracT2
CW_etmiss_mean	1.000	0.053	0.016	-0.017	-0.000	0.277	-0.128
CW_mjjj_base	0.053	1.000	-0.091	0.236	-0.000	0.121	-0.253
G_etmiss_mean	0.016	-0.091	1.000	-0.050	-0.000	0.093	0.158
G_mjjj_sigma	-0.017	0.236	-0.050	1.000	-0.000	-0.046	-0.075
G_mtrans_mean	-0.000	-0.000	-0.000	-0.000	1.000	-0.000	0.000
SM_fracCW	0.277	0.121	0.093	-0.046	-0.000	1.000	-0.294
TT_fracT2	-0.128	-0.253	0.158	-0.075	0.000	-0.294	1.000

Table 7.19: The top half of this table shows the fit result of the background model to data in the control region. The bottom half shows the extrapolated background yields in the signal region. The `mean` and `sigma` parameters are in units of GeV, the `base` parameters in units of GeV^{-1}.

7.5.2 Fitting the model with SUSY Ansatz to data

Although the fit result of the background-only model seems quite consistent with the expectation from MC, it is nevertheless interesting to see if the complete model finds some evidence of SUSY contamination in the control region, which might lead to an over-estimation of the background contribution in the signal region. Table 7.20 shows the result of this fit. The SM background yields are compatible with MC expectations. A small SU yield is fitted, which is however consistent with 0 SUSY events in the control region. The expected number of background events in the signal region is compatible with the measurement (1.8 ± 0.9(stat) events from the extrapolation of the fit in the CR, vs 1 measured).

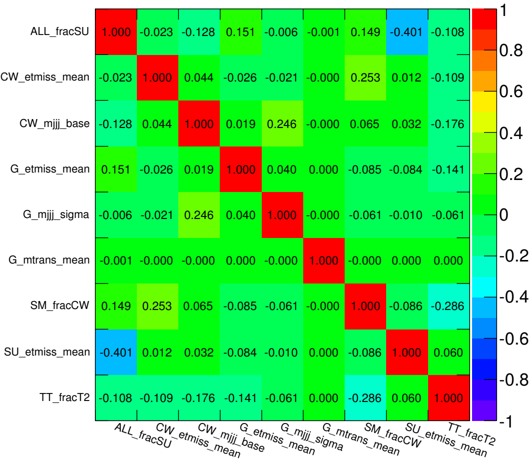
7.5.3 Systematics

Two distinct types of systematic uncertainties will be discussed in this section. First, there are systematic uncertainties directly related to the measurement by the detector. These are the jet energy scale (JES), jet momentum resolution, $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, lepton energy scale and lepton momentum resolution. These uncertainties, which are common to all ATLAS analyses using these observables, will be discussed in the first part of this section.

The second type of systematic errors are specific to this analysis. They are induced by the shape parameters that are fixed by the MC simulation. To reduce this uncertainty, those parameters that induce the largest systematic uncertainty have been allowed to float in the combined fit. Related to this type of error is the systematic error induced by choosing certain MC generators. To estimate the size of this error an ALPGEN $t\bar{t}$ sample was used.

The systematic uncertainties quoted in this section are uncertainties on the estimated

Sample	Fitted yield	MC expectation
Control region		
TP	7.2 ± 3.7	5.3
CW	30 ± 5.2	30.5
T2	4.1 ± 3.9	2.5
SU	2.2 ± 2.4	–
Signal region		
TP	0.3 ± 0.2	0.05
CW	0.8 ± 0.2	0.3
T2	0.8 ± 0.7	0.3
Fitted shape parameters		
`CW_etmiss_mean`		67 ± 4
`CW_mjjj_base`		$-(10 \pm 2) \cdot 10^{-3}$
`G_etmiss_mean`		104 ± 10
`G_mjjj_sigma`		41 ± 11
`G_mtrans_mean`		100 ± 3
`SU_etmiss_mean`		238 ± 48



	ALL_fracSU	CW_etmiss_mean	CW_mjjj_base	G_etmiss_mean	G_mjjj_sigma	G_mtrans_mean	SM_fracCW	SU_etmiss_mean	TT_fracT2
ALL_fracSU	1.000	-0.023	-0.128	0.151	-0.006	-0.001	0.149	-0.401	-0.108
CW_etmiss_mean	-0.023	1.000	0.044	-0.026	-0.021	-0.000	0.253	0.012	-0.109
CW_mjjj_base	-0.128	0.044	1.000	0.019	0.246	-0.000	0.065	0.032	-0.176
G_etmiss_mean	0.151	-0.026	0.019	1.000	0.040	0.000	-0.085	-0.084	-0.141
G_mjjj_sigma	-0.006	-0.021	0.246	0.040	1.000	-0.000	-0.061	-0.010	-0.061
G_mtrans_mean	-0.001	-0.000	-0.000	0.000	-0.000	1.000	-0.000	0.000	0.000
SM_fracCW	0.149	0.253	0.065	-0.085	-0.061	-0.000	1.000	-0.086	-0.286
SU_etmiss_mean	-0.401	0.012	0.032	-0.084	-0.010	0.000	-0.086	1.000	0.060
TT_fracT2	-0.108	-0.109	-0.176	-0.141	-0.061	0.000	-0.286	0.060	1.000

Table 7.20: The top half of this table shows the fit result of the full (background + SUSY Ansatz) model to data in the control region. The bottom half shows the extrapolated background yields in the signal region. The `mean` and `sigma` parameters are in units of GeV, the `base` parameters in units of GeV^{-1}.

number of background events in the signal region.

7.5.4 Detector systematics

The detector systematic uncertainties are listed in Table 7.22. Each energy scale and resolution variation affects not only the object in question (i.e. jets for the jet energy scale (JES), muon for the muon p_T resolution), but also the value of $E_{\mathrm{T}}^{\mathrm{miss}}$, as it is calculated from all objects in the event. Each source of uncertainty was scaled (up and down) by one standard deviation, and all relevant quantities were recalculated. See Table 7.21 for the size of the uncertainties of the quantities that were taken into account. Then the dataset was used as input to the combined fit method, and the expected background yield in the signal region was recalculated. Table 7.22 quotes the relative change (averaged for scaling up or down) in the expected background yield with regards to the original estimate.

Clearly the jet energy scale has the biggest effect on the outcome of the fit. This has two reasons. It is expected to be a large systematic, since the JES has a correlated effect on all jets in the event (unlike jet energy resolution, where the effect is uncorrelated), and events are selected with high jet multiplicity. The effect of the JES is further increased by the fact that also the $E_{\mathrm{T}}^{\mathrm{miss}}$ value is corrected when the jet energies are corrected, so the JES has a large effect both on the event selection efficiency and on the shape of all three distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj}.

Source	error
JES barrel	5%
JES endcap	7%
p_T^{μ} scale	1%
p_T^{jet} resolution	10%
p_T^{μ} resolution (barrel)	5%
p_T^{μ} resolution (endcap)	8%
$E_{\mathrm{T}}^{\mathrm{miss}}$ resolution	10%

Table 7.21: The uncertainties on different observables, used as input for the systematic uncertainty analysis.

Source	systematic error
JES	51 %
μ p_T scale	1 %
Jet p_T resolution	9 %
μ p_T resolution	2 %
$E_{\mathrm{T}}^{\mathrm{miss}}$ resolution	9 %
Detector systematic error	52%

Table 7.22: Sources and size of detector systematic errors.

7.5.5 Shape systematics

Due to the limited size of the data sample the combined fit background model and the SUSY Ansatz do not allow for all shape parameters to be determined in the fit. Some parameters need to be fixed after the prefit. For the background model, the decision which parameters were fixed was based on Table 7.23. This table was generated by performing a baseline fit of the complete model on the data measured in the control region, with each parameter at the prefit value. Then for each parameter, the same fit was done with the parameter value scaled up and down by one standard deviation. The relative shift of the extrapolated SM yield in the signal region (again averaged for up or down scaling) is quoted in the table as the systematic error.

Which parameter was released first was determined by the size of the associated systematic error in Table 7.23. It turned out that most parameters that were not part of the $T2$ model could be released. Table 7.24 shows the remaining systematic error from fixed shape parameters with as many parameters floating as possible. This table also shows the error induced by the choice of MC generators. To estimate this error an ALPGEN $t\bar{t}$ sample was used in the prefit stage, and the resulting model was fitted to a simulated sample with MC@NLO $t\bar{t}$ (and the original W+jets sample). Note that the systematic error with all shapes fixed (in Table 7.23) is significantly larger than when some of the shape parameters are floating in the fit (Table 7.24). Note furthermore that as the size of the data sample increases, the fit will become more stable, and the number of floating parameters may increase, thus reducing the systematic error even more.

In building the model several parameters of the model were shared if they had comparable

Shape parameter	systematic error
`G_mtrans_base`	102 %
`G_mtrans_mean`	91 %
`T2_mtrans_base`	64 %
`T2_mtrans_sigma`	50 %
`T2_mtrans_mean`	48 %
MC@NLO versus ALPGEN	43 %
`G_mtrans_sigma`	35 %
`G_etmiss_mean`	25 %
`T2_mjjj_base`	25 %
`G_etmiss_base`	19 %
`CW_etmiss_mean`	11 %
`T2_etmiss_sigma`	9 %
`CW_etmiss_sigma`	7 %
`CW_mjjj_base`	6 %
`G_mjjj_mean`	4 %
`TP_mjjj_mean`	3 %
`TP_etmiss_sigma`	2 %
`TP_mjjj_sigma`	2 %
`G_mjjj_sigma`	0.7 %

Table 7.23: For each parameter of the model, this table shows the estimate of the systematic error associated with keeping that parameter fixed. Based on this table, some parameters have been released in the fit. The ALPGEN systematic error was estimated by prefitting the shape parameters on an ALGPEN $t\bar{t}$ sample, instead of MC@NLO.

value (global parameters). For instance, the position of the M_{T} peak is expected to be the same for both the TP and the CW model, as it is determined by the W mass in both cases. To estimate the effect on the outcome of the fit the model was fitted with fixed shape parameters, once with global parameters, and once without. The relative difference is also quoted in Table 7.24.

For the SUSY Ansatz, the assumption was made that the correlation between the Ansatz shape parameters could be fixed. Table 7.25, which was obtained using the same procedure as described above, shows that the systematic error induced by this assumption is small for each correlation parameter.

7.5.6 Limits on mSUGRA phase space

The SM background estimate from section 7.5.2 can be used to set a limit on the mSUGRA phase space. This limit was set by doing a simple counting experiment in the signal region, taking the background estimate and the number of observed events as input to a Poisson distribution:

$$L(N_{obs}) = Poisson(s + b) \tag{7.5}$$

with b the expected mean number of background events, and s the mean number of signal events. The systematic uncertainties on s and b and the uncertainty from the control region

Shape parameter	systematic error
`T2_mtrans_base`	49 %
`T2_mtrans_sigma`	38 %
`T2_mtrans_mean`	35 %
`T2_etmiss_sigma`	19 %
`T2_mjjj_base`	13 %
`G_mtrans_sigma`	8 %
`TP_etmiss_sigma`	8 %
`G_etmiss_base`	6 %
`TP_mjjj_mean`	4 %
`TP_mjjj_sigma`	4 %
`G_mtrans_base`	14 %
`G_mtrans_sigma`	8 %
`CW_etmiss_sigma`	8 %
`G_mjjj_mean`	2 %
MC@NLO versus ALPGEN	18 %
Global shape parameters	8%
Total shape systematic error	80%

Table 7.24: Sources and size of systematic uncertainty of the constant shape parameters errors. This table lists all the parameters that were kept constant during the fit procedure. The first ten parameters were kept fixed because releasing them would cause fit instability. The last four parameters were kept fixed because of large correlations with other parameters. The ALPGEN systematic error was estimated by prefitting the shape parameters on an ALPGEN $t\bar{t}$ sample, instead of MC@NLO. The systematic error for global shape parameters represents the relative difference of a fit to data with all shape parameters fixed, with or without sharing parameters between two models.

Shape parameter	systematic error
`SU_mjjj_mean_offset`	0.1 %
`SU_mjjj_mean_slope`	0.1 %
`SU_mjjj_sigma_offset`	0.1 %
`SU_mjjj_sigma_slope`	0.1 %
`SU_mtrans_base_offset`	0.1 %
`SU_mtrans_base_slope`	0.1 %
`SU_mjjj_base_offset`	0.09 %
`SU_mjjj_base_slope`	0.09 %
`SU_etmiss_sigma_slope`	0.08 %
`SU_etmiss_sigma_offset`	0.05 %

Table 7.25: Sources and size of shape systematic of the SUSY Ansatz shape parameters errors.

fit are treated as nuisance parameters in this likelihood, thus giving:

$$L(N_{obs}) = \sum_i e^{(1-r_b^i)^2/\sigma_i^2} \sum_j e^{(1-r_s^j)^2/\sigma_j^2} Poisson(r_s s + r_b b) \tag{7.6}$$

with σ_i and σ_j the uncertainties on the signal and background expectations, respectively, $r_{s,b} = \prod_i r_{s,b}^i$ are a product of all corrections on s and b induced by nuisance parameters.

All the systematic errors of Table 7.22 were taken into account. The systematic error induced by the background shape parameters was estimated by taking the quadratic sum of all errors in Table 7.23. The systematic error induced by the SUSY shape parameters was so small it was ignored.

Two systematic errors were taken into account that affect the signal parameter s: the 3.4% uncertainty on the luminosity, and an average theoretical uncertainty of 20% on the SUSY cross section [68], which was assumed to be identical for all points.

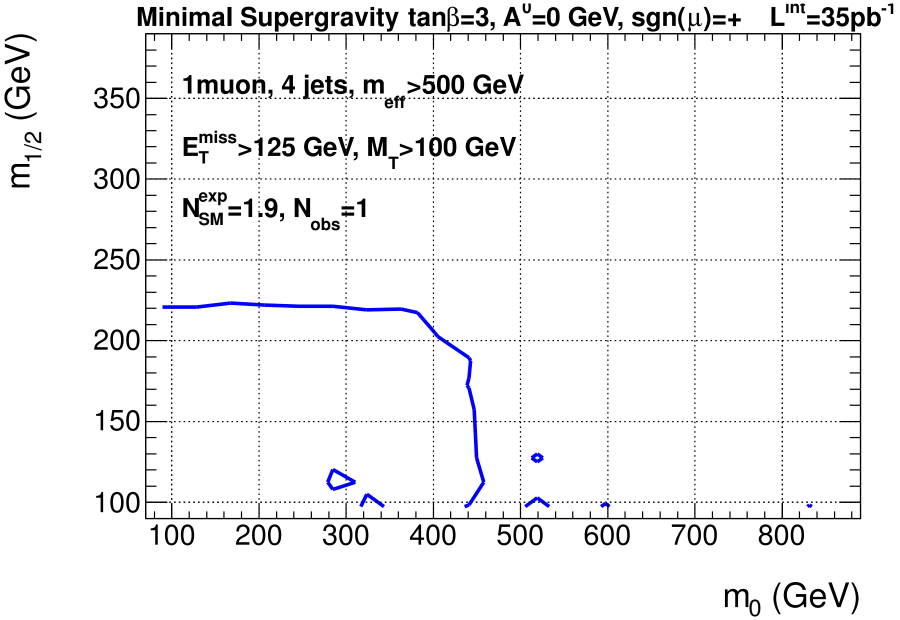


Figure 7.9: The limit in the $m_0/m_{1/2}$ plane set using the combined fit method.

According to Wilks theorem [69] the profile likelihood ratio can be used as a test statistic, which follows a χ^2 distribution for the given number of the degrees of freedom. This can then be used to calculate the 95% confidence limit on the number of signal events in the signal region. It must be noted that Wilks theorem is in principle only valid for a large number of measurements, but it has been shown by members of the ATLAS SUSY group [70] that the difference between the profile likelihood and a full toy Monte Carlo analysis is small. It does mean that the limit set in this section is slightly too aggressive.

Applying this procedure to compare the Combined fit method prediction to the ATLAS measurement, a limit can be set on the number of signal events in the signal region of 3.5 at 95 % confidence level.

The limit on the number of signal events present in the signal region was compared to the expectation for all points in the mSUGRA grid. This resulted in the limit in mSUGRA phase space set by the combined fit method, which is shown in Figure 7.9. This contour can be translated into squark/gluino masses as follows:

- $m_0 = 200\,\text{GeV}$, $m_{1/2} = 220\,\text{GeV}$: $m_{\tilde{g}} = 545\,\text{GeV}$, $m_{\tilde{q}} = 530\,\text{GeV}$,
- $m_0 = 360\,\text{GeV}$, $m_{1/2} = 220\,\text{GeV}$: $m_{\tilde{g}} = 560\,\text{GeV}$, $m_{\tilde{q}} = 600\,\text{GeV}$,
- $m_0 = 440\,\text{GeV}$, $m_{1/2} = 185\,\text{GeV}$: $m_{\tilde{g}} = 490\,\text{GeV}$, $m_{\tilde{q}} = 600\,\text{GeV}$,
- $m_0 = 450\,\text{GeV}$, $m_{1/2} = 150\,\text{GeV}$: $m_{\tilde{g}} = 400\,\text{GeV}$, $m_{\tilde{q}} = 550\,\text{GeV}$.

These limits are a significant improvement on the previous limit set by the D∅ experiment of $m_{\tilde{g}} = 308\,\text{GeV}$ and $m_{\tilde{q}} = 379\,\text{GeV}$ [22], and of the CDF experiment of $m_{\tilde{g}} = 350\,\text{GeV}$ [71].

Conclusion and Outlook

Conclusion

This thesis describes the Combined Fit Method, providing a data driven method to estimate the Standard Model backgrounds to supersymmetry. These backgrounds are estimated using a 3D model describing the distributions of the three observables $E_{\mathrm{T}}^{\mathrm{miss}}$, M_{T} and M_{jjj}. This model with parametrized shape is fitted to data in a control region with low $E_{\mathrm{T}}^{\mathrm{miss}}$ or low M_{T}, thus minimizing the dependance on Monte Carlo simulation. This fit takes contamination of mSUGRA like signatures into account, preventing the overestimation of the SM background in the signal region, thus increasing the sensitivity of the method to new physics.

The analysis on Monte Carlo simulated samples demonstrated the power of the method on a dataset of about $1\,\mathrm{fb}^{-1}$, with a center of mass energy of the pp collisions of $10\,\mathrm{TeV}$, correlations between the three observables cannot be ignored. Using conditional PDFs these correlations have been incorporated in the model. The mathematics of dealing with these conditional PDFs in the context of this analysis has been investigated in this thesis, and the necessary formulas have been incorporated in the `RooFit` framework.

The $10\,\mathrm{TeV}$ analysis ended with a proof of principle on a Monte Carlo simulated pseudodata sample, demonstrating that the method is capable of correctly estimating the shape of the background contributions, also when the shape of the pseudodata is different than the input shapes of the model. Specifically typical correlations between the observables helped to fit the dileptonic $t\bar{t}$ background, the most SUSY like of all SM backgrounds. The yield of mSUGRA events in the pseudodata was correctly estimated by the method in both control and signal region, proving that the SUSY Ansatz worked well. A small positive bias was found when performing a toy Monte Carlo analysis, which can be explained by an asymmetry in the distribution of mSUGRA events when the expected number of events is small.

With an integrated luminosity of $35\,\mathrm{pb}^{-1}$ and a center of mass energy of the pp collision of $7\,\mathrm{TeV}$, the 2010 dataset is much smaller than the simulated sample the model was created and tested on. Larger statistical errors meant on the one hand that the model could be much simplified, and on the other that it would become more difficult to distinguish SM background from SUSY signal. Furthermore, the part of mSUGRA phase space that the analysis on data is sensitive to produces overall a much more SM-like signature than the mSUGRA models studied in the MC analysis.

The model was simplified in two ways. First, the functional form used to describe the shapes of the distributions is much simpler, and thus contains much less shape parameters than before. For instance instead of describing the M_{T} distribution with the MTFunc PDF with five parameters, the TTComb with only three parameters suffices. Second, the correlations between the observables were demonstrated to be small compared to the statistical uncertainty, and could be completely ignored. The SUSY Ansatz model was also simplified, and correlations between shape parameters allowed the model to be described with a single shape parameter. The SUSY Ansatz helps to reduce the overestimation of the SM in the control region, and hence also the extrapolated SM expectation. The estimate of the SUSY contamination gets worse however as the SUSY mass scale gets smaller, as the SUSY signal becomes more and more SM like. These low mass SUSY points will however be noticed in the data due to a very high SM cross section. No such large excess of SM-like events has been seen in data.

The analysis of the systematic error associated with the shapes of the distributions revealed an important advantage of using this data driven method. These errors can be as large as 100%. Using the ALPGEN generator for the $t\bar{t}$ sample to estimate the initial values of the shapes revealed a systematic error based on generator choice as large as 40%, when all the shape parameters were fixed.

By estimating the value of many of the shape parameters in a fit to data the largest systematic error induced by a shape parameter was reduced to about 50%, caused by the dileptonic shape parameters which were kept at a fixed value to avoid confusion between this background and SUSY. The systematic error induced by generator choice is now only 7%. Fixing the correlations in the SUSY Ansatz model was shown to lead to a systematic error of less than 1%.

The combined fit method predicted 1.8 ± 0.9 events to be measured in the signal region. This means the observed number of events (1 in the signal region) is compatible with the SM expectation. Using a profile likelihood ratio, a limit on the possible SUSY yield in the signal region of 3.5 events at 95% CL. This limit has been compared to the expected yields of a grid of points in mSUGRA phase space shown in figure 7.9.

The analysis of the ATLAS SUSY group [67] found an upper limit of 2.5 signal events in the signal region. That analysis differs from the one in this thesis in a number of respects. Somewhat different event selection cuts were applied. For a prediction of the SM background in the signal region, the ATLAS analysis used the number of events in dedicated control regions as normalization, but relied on Monte Carlo to derive a prediction from this in the signal region; in other words Monte Carlo shapes were used. Of course a systematic uncertainty was assigned to this procedure, but the procedure in this thesis is more data-driven. The slightly lower limit does not significantly improve the area of the mSUGRA phase space that can be excluded. The statistical uncertainty on the estimate will however rapidly decrease with increasing size of the data sample. This means that the systematic uncertainty caused by relying on MC to predict the shape of the distributions will become more and more important. Plans are therefore now being made in the ATLAS SUSY group to incorporate the Combined Fit method in their standard analysis.

Outlook

The statistical errors on the number of events still plays a large role in limiting the power of the combined fit method. Not only will the estimate of the fitted yields become more accurate with increasing statistics, also more details such as the correlations in the dileptonic $t\bar{t}$ distributions will become visible. As chapter 6 has shown this method is capable of dealing with this details.

When more events are available, the method can also be expanded. More control regions can be defined where specific backgrounds are enhanced. By using the `RooFit` feature of fitting the model simultaneously in multiple control regions, the shapes of the background model can be estimated with much better accuracy.

One piece of information that could be used in defining extra control regions is b-tagging. Only $t\bar{t}$ events are expected to produce significant number of b-jets, thus splitting the data sample in events with 0, 1 or 2 b-tags can provide a control sample with much enhanced $t\bar{t}$ yield.

As SUSY events are characterized by a large number of jets, adding events with only 3 jets will provide a sample that contains more background relative to the number of SUSY events. This sample can be used to estimate some of the parameters of the background model for the 4-jet sample. Specifically the M_T distribution is not expected to change when looking at a different number of jets.

Since the parton distribution of protons has an overall positive charge, a sample of W+jets events contain more W^+ then W^-. This charge asymmetry is not present in pair produced $t\bar{t}$ or SUSY events. By measuring the charge of the lepton, two samples can be created where the yield of W events is different in each sample, giving an extra constraint when estimating the shape parameters. The fitted asymmetry can furthermore be used as a cross check which can be compared with the W analysis of ATLAS. Splitting the top sample and combining W with top combinatorics will however complicate the use of charge asymmetry.

Another interesting topic would be using a broader Ansatz model, that is valid for different types of (supersymmetric) theories beyond the Standard Model. As the typical mass scale of the new theory seems to be the main factor in determining the behavior for low values of $E_\mathrm{T}^\mathrm{miss}$ and M_T, it is conceivable that the application of the Ansatz could be broadened.

Finally the Combined Fit Method can be used in the analysis of the ATLAS SUSY group, in order to reduce the systematic uncertainty associated with the use of Monte Carlo, and to increase the robustness of the ATLAS supersymmetry search.

Bibliography

[1] S. L. Glashow. Partial-symmetries of weak interactions . *Nuclear Physics*, 22(4):579–588, Feb 1961.

[2] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19(21):1264–1266, Nov 1967.

[3] A. Salam and J. C. Ward. Weak and electromagnetic interactions. *Il Nuovo Cimento*, 11:568–577, 1959.

[4] J. Goldstone. Field theories with superconductor solutions. *Il Nuovo Cimento (1955-1965)*, pages 154–164, 1961.

[5] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13(16):508–509, Oct 1964.

[6] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13(20):585–587, Nov 1964.

[7] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13(9):321–323, Aug 1964.

[8] W. J. Stirling R. K. Ellis and B. R. Webber. *QCD and Collider Physics.* Cambridge University Press, 1996.

[9] D. V. Schroeder M. E. Peskin. *An Introduction to Quantum Field Theory.* Westview Press, 1995.

[10] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tonnesmann. Jets in hadron-hadron collisions. *Prog. Part. Nucl. Phys.*, 60:484–551, 2008, hep-ph/0712.2447.

[11] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002, hep-ph/0201195.

[12] K. Nakamura et al. (Particle Data Group). 011 review of particle physics. *J. Phys.*, G 37, 2010.

[13] S. P. Martin. A supersymmetry primer, 2008, hep-ph/9709356v5.

[14] I. J. R. Aitchison. Supersymmetry and the MSSM: An Elementary Introduction, 2005, hep-ph/0505105v1.

[15] H. Flacher et al. Gfitter - Revisiting the Global Electroweak Fit of the Standard Model and Beyond. *Eur.Phys.J.*, C60:543–583, 2009, hep-ph/0811.0009.

[16] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys.Lett.*, B565:61–75, 2003, hep-ex/0306033.

[17] Y. Kazuhiro. Higgs searches at the Tevatron. *Int.J.Mod.Phys.*, A25:5097–5104, 2010.

[18] D. Ludwig. The global electroweak fit and constraints on new physics with Gfitter. *PoS*, ICHEP2010:404, 2010, hep-ph/1010.5678.

[19] N. Palanque-Delabrouille. Overview of astroparticle physics and dark matter searches. *Int.J.Mod.Phys.*, A22:5735–5746, 2007.

[20] G. W. Bennett et al. Measurement of the negative muon anomalous magnetic moment to 0.7 ppm. *Phys.Rev.Lett.*, 92:161802, 2004, hep-ex/0401008.

[21] J. R. Ellis, J. S. Hagelin, and D. V. Nanopoulos. Spin 0 leptons and the anomalous magnetic moment of the muon. *Phys.Lett.*, B116:283, 1982.

[22] V.M. Abazov et al. Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96$- TeV. *Phys.Lett.*, B660:449–457, 2008, hep-ex/0712.3805.

[23] L. Evans, (ed.) and P. Bryant, (ed.). LHC Machine. *JINST*, 3:S08001, 2008.

[24] The ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[25] R. Adolphi et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[26] A. A. Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[27] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[28] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008.

[29] O. Adriani et al. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008.

[30] LCG collaboration. LCG website. `http://lcg.web.cern.ch/`.

[31] The ATLAS collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. Technical report, CERN, 2009, hep-ex/0901.0512.

[32] C. Schwanenberger. The jet calibration in the H1 liquid argon calorimeter. Technical report, Desy, 2002, physics/0209026.

[33] M. Cacciari and G. P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett.*, B641:57–61, 2006, hep-ph/0512210.

[34] The ATLAS collaboration. Observation of inclusive electrons in the ATLAS experiment at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-073, CERN, Geneva, Jul 2010.

[35] W. Verkerke and I. van Vulpen. Commissioning atlas using top-quark pair production. Technical Report ATL-COM-PHYS-2007-023, CERN, Geneva, Apr 2007.

[36] The ATLAS collaboration. ATLAS public results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/WebHome`.

[37] The ATLAS collaboration. Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the atlas experiment. Technical Report ATLAS-CONF-2010-069, CERN, Geneva, Jul 2010.

[38] The ATLAS collaboration. J/ψ Performance of the ATLAS Inner Detector. Technical Report ATLAS-CONF-2010-078, CERN, Geneva, Jul 2010.

[39] The ATLAS collaboration. Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2011-012, CERN, Geneva, Mar 2011.

[40] The ATLAS collaboration. Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the atlas detector using proton-proton collisions at a center of mass energy $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-054, CERN, Geneva, Jul 2010.

[41] The ATLAS collaboration. Reconstruction and calibration of missing transverse energy and performance in z and w events in atlas proton-proton collisions at 7 tev. Technical Report ATLAS-CONF-2011-080, CERN, Geneva, Jun 2011.

[42] The ATLAS collaboration. Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS 2010 data. Technical Report ATLAS-CONF-2011-032, CERN, Geneva, Mar 2011.

[43] The ATLAS collaboration. Muon reconstruction efficiency in reprocessed 2010 lhc proton-proton collision data recorded with the atlas detector. Technical Report ATLAS-CONF-2011-063, CERN, Geneva, Apr 2011.

[44] The ATLAS collaboration. Muon momentum resolution in first pass reconstruction of pp collision data recorded by atlas in 2010. Technical Report ATLAS-CONF-2011-046, CERN, Geneva, Mar 2011.

[45] The ATLAS collaboration. Performance of the atlas muon trigger in p-p collisions at $\sqrt{s}$ = 7 TeV. Technical Report ATLAS-CONF-2010-095, CERN, Geneva, Oct 2010.

[46] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling, 2003, physics/0306116.

[47] R. Brun and F. Rademakers. ROOT: An object oriented data analysis work. *Nucl. Instrum. Meth.*, A389:81–86, 1997.

[48] F. Koetsveld and A. J. Koutsman. The combined fit method for 1-lepton Susy searches. Technical report, CERN, 2010. ATL-COM-PHYS-2010-899.

[49] A. J. Koutsman. *On the road to Supersymmetry with ATLAS*. PhD thesis, Radboud University Nijmegen, 2011.

[50] The ATLAS Computing Group. ATLAS Computing technical design report. Technical report, CERN, 2005. CERN-LHCC-2005-022.

[51] K. Assamagan et al. The ATLAS Monte Carlo Project. Technical Report ATL-SOFT-INT-2010-002, CERN, Geneva, Feb 2010.

[52] S. Agostinelli et al. GEANT4 - A Simulation Toolkit. *Nuclear Instruments and Methods*, A506:250–303, 2003.

[53] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 06:029, 2002, hep-ph/0204244.

[54] S. Frixione, P. Nason, and B. R. Webber. Matching NLO QCD and parton showers in heavy flavour production. *JHEP*, 08:007, 2003, hep-ph/0305252.

[55] G. Corcella et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001, hep-ph/0011363.

[56] G. Corcella et al. HERWIG 6.5 release note, 2002, hep-ph/0210213.

[57] A. Shibata et al. Understanding monte carlo generators for top physics. Technical Report ATL-COM-PHYS-2009-334, CERN, Geneva, Jun 2009.

[58] M. L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 07:001, 2003, hep-ph/0206293.

[59] S. Hoeche et al. Matching parton showers and matrix elements, 2006, hep-ph/0602031.

[60] F. Ahles et al. Prospects for Susy and UED discovery based on inclusive searches at 10 TeV centre-of-mass energy with the ATLAS detector. Technical Report ATL-PHYS-INT-2009-060. ATL-COM-PHYS-2009-302, CERN, Geneva, May 2009.

[61] F. E. Paige et al. ISAJET 7.69: A Monte Carlo event generator for p p, anti-p p, and e+ e- reactions, 2003, hep-ph/0312045.

[62] L. Moneta et al. The RooStats Project, 2010, physics.data-an/1009.1003v1.

[63] Cranmer, K. Statistical Challenges for Searches for New Physics at the LHC. In L. Lyons & M. Karagöz Ünel, editor, *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, 2006, physics/0511028.

[64] R.J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* Wiley, July 1989.

[65] W. Verkerke. Data analysis lectures. Amsterdam, 2004. `http://www.nikhef.nl/~verkerke/talks/data_analysis/data_analysis_2004_v17.pdf`.

[66] The ATLAS collaboration. Updated luminosity determination in pp collisions at $\sqrt{s} = 7$ tev using the atlas detector. Technical Report ATLAS-CONF-2011-011, CERN, Geneva, Mar 2011.

[67] The ATLAS collaboration. Search for supersymmetry using final states with one lepton, jets, and missing transverse momentum with the ATLAS detector in $\sqrt{s} = 7$ TeV pp. *Phys.Rev.Lett.*, 106:131802, 2011, 1102.2357.

[68] R. H. L. van der Leeuw, O. Igonkina, and P. de Jong. Systematic Uncertainties in mSUGRA NLO cross-section calculations at $\sqrt{s} = 7$ TeV. Technical Report ATL-PHYS-INT-2010-086, CERN, Geneva, Sep 2010.

[69] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Ann. Math. Statist.*, 9(1):60–62, 1938.

[70] M. Baak. private communication, 2001.

[71] T. Aaltonen et al. Search for Gluino-Mediated Sbottom Production in $p\bar{p}$ Collisions at $\sqrt(s) = 1.96$ TeV. *Phys. Rev. Lett.*, 102:221801, 2009, hep-ex/0903.2618.

Appendix A

Combined Fit Model on simulation

The combined fit model

The combined fit model looks like:

$$TT = (1 - f_{T2})TP + f_{T2}T2 \tag{A.1}$$

$$SM = (1 - f_{TT})CW + f_{TT}TT \tag{A.2}$$

$$ALL_{S+B} = (1 - f_{SM})SU + f_{SM}SM \tag{A.3}$$

where TP denotes the PDF describing the TP background etcetera. Each component model of ALL_{S+B} is a 3D product PDF. Each background model is a 3D conditional product PDF. The composition of these PDFs is given in table A.1 t/m A.4. Each table first lists the structure of the models and their parameters, where functions are denoted in boldface. Then the functions that are used inside a model, i.e. the conditional parameters that are a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the interdependencies of the parameters of the SU model are listed.

The naming of PDFS and their parameters is as follows. A 3D PDF is named after the background source it describes. A 1D PDF name contains the source and the observable like < `source`>_< `observable`>. If a PDF is a sum of more than one component, the component PDF names are written as < `source`>_< `observable`>_< `suffix`>, where < `suffix`> will denote the components function like `_peak` or `_tail`. A parameter of a PDF is denoted as < `source`>_< `observable`>_< `parameter`>. If this parameter is a function of $E_{\mathrm{T}}^{\mathrm{miss}}$ or another parameter, the name of the parameters of this function is denoted as < `source`>_< `observable`>_< `function`>_< `parameter`>. This system ensures unique names for all PDFs, functions and parameters.

PDFs		
PDF	Component	Parameters
`TTComb:TP_etmiss`		`TP_etmiss_base`, `G_etmiss_mean`, `TP_etmiss_sigma`
`MTFunc:TP_mtrans`		`TP_mtrans_mean`, `G_mtrans_sigma`, **TP_mtrans_mean_ratio**, `TP_mtrans_sigma_ratio`, **TP_mtrans_fpeak**, `TP_mtrans_base`
`AddPdf:TP_mjjj`	`Gaussian:TP_mjjj_peak`	`TP_mjjj_peak_mean`, `TP_mjjj_peak_sigma`
	`TTComb:TP_mjjj_comb`	`TP_mjjj_base`, `TP_mjjj_comb_mean`, `TP_mjjj_comb_sigma`

Functions	
Function name	Function form
TP_mtrans_mean_ratio	max(0.5,(G_mtrans_mean_ratio_offset + G_mtrans_mean_ratio_slope×$E_{\mathrm{T}}^{\mathrm{miss}}$))
TP_mtrans_fpeak	min(1,(TP_mtrans_fpeak_offset - G_mtrans_fpeak_amp × **TP_mtrans_fpeak_erf**))
TP_mtrans_fpeak_erf	Erf((etmiss-TP_mtrans_fpeak_mean)/ TP_mtrans_fpeak_sigma)

Table A.1: The TP model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

PDFs

PDF	Component	Parameters
TTComb:T2_etmiss		T2_etmiss_base, G_etmiss_mean, T2_etmiss_sigma
MTFunc:T2_mtrans		T2_mtrans_corefrac **T2_mtrans_mean**, **T2_mtrans_sigma**, T2_mtrans_mean_ratio, T2_mtrans_sigma_ratio, **T2_mtrans_fpeak**, **T2_mtrans_base**
TTComb:T2_mjjj		**T2_mjjj_base**, **T2_mjjj_mean**, G_mjjj_sigma

Functions

Function name	Functional form
T2_mtrans_mean	min(250,(T2_mtrans_mean_slope × etmiss + T2_mtrans_mean_offset))
T2_mtrans_sigma	min(35,(T2_mtrans_sigma_slope × etmiss + T2_mtrans_sigma_offset))
T2_mtrans_fpeak	min(1.,(T2_mtrans_fpeak_offset + T2_mtrans_fpeak_amp × **T2_mtrans_fpeak_erfc**))
T2_mtrans_base	min(-0.0001,(T2_mtrans_base_slope × etmiss + T2_mtrans_base_offset + T2_mtrans_base_amp × **T2_mtrans_base_erf**))
T2_mtrans_fpeak_erfc	Erfc((etmiss-T2_mtrans_fpeak_mean) / T2_mtrans_fpeak_sig)
T2_mtrans_base_erf	Erf((etmiss-T2_mtrans_base_mean) / T2_mtrans_base_sig)

Table A.2: The $T2$ model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

PDFs

PDF	Component	Parameters
AddPdf:CW_etmiss	TTComb:CW_etmiss1	CW_etmiss_base, CW_etmiss_mean, CW_etmiss_sigma
	TTComb:CW_etmiss2	**CW_etmiss_base2**, **CW_etmiss_mean2**, CW_etmiss_sigma
MTFunc:CW_mtrans		G_mtrans_corefrac, **CW_mtrans_mean**, G_mtrans_sigma, **CW_mtrans_mean_ratio**, G_mtrans_sigma_ratio, **CW_mtrans_fpeak**, CW_mtrans_base
TTComb:CW_mjjj		**CW_mjjj_base**, **CW_mjjj_mean**, G_mjjj_sigma

Functions

Function name	Functional form
CW_etmiss_base2	CW_etmiss_base_ratio × CW_etmiss_base
CW_etmiss_mean2	CW_etmiss_mean_ratio × CW_etmiss_mean
CW_mtrans_mean	CW_mtrans_mean_slope × etmiss + CW_mtrans_mean_offset
CW_mtrans_mean_ratio	G_mtrans_mean_ratio_slope × etmiss + G_mtrans_mean_ratio_offset
CW_mtrans_fpeak	G_mtrans_fpeak_offset - G_mtrans_fpeak_amp × **CW_mtrans_fpeak_erf**))
CW_mjjj_base	CW_mjjj_base_offset + CW_mjjj_base_amp × **CW_mjjj_base_erf**
CW_mjjj_mean	CW_mjjj_mean_slope × etmiss + CW_mjjj_mean_offset
CW_mtrans_fpeak_erf	Erf((etmiss-CW_mtrans_fpeak_mean)/ CW_mtrans_fpeak_sig)
CW_mjjj_base_erf	Erf((etmiss-CW_mjjj_base_mean)/ CW_mjjj_base_sigma)

Table A.3: The *CW* model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

	PDFs	
PDF	Component	Parameters
TTComb:SU_etmiss		SU_etmiss_base, **SU_etmiss_mean**, **SU_etmiss_sigma**
AddPdf:SU_mtrans	Uniform:SU_mtrans_1	
	TTComb:SU_mtrans_real	SU_mtrans_base, **SU_mtrans_mean**, SU_mtrans_sigma
Decay:SU_mjjj		**SU_mjjj_mean**, **SU_mjjj_tau**, **SU_mjjj_sigma**
	Functions	
Function name	Functional form	
SU_etmiss_mean	$2.4\times$ SU_etmiss_sigma-16.5	
SU_etmiss_sigma	$7.7\cdot 10^3\times$ SU_etmiss_base$+157$	
SU_mtrans_mean	$-7.8\cdot 10^3\times$ SU_mtrans_base-52.6	
SU_mjjj_mean	$0.43\times$ SU_mjjj_tau$+39.6$	
SU_mjjj_tau	$1.8\cdot 10^4\times$ SU_etmiss_base$+436$	
SU_mjjj_sigma	$0.54\times$ SU_mjjj_mean-37.7	

Table A.4: The *SU* model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

Parameters in maximum floating shapes model

The complete model as described above contains 74 distinct parameters. This is more then the minimum amount of parameters necessary to describe the signal and background of this analysis. This is clearly visible through large correlations between parameters of the model. Only 29 shape parameters + 3 yield parameters are 'real' parameters of the fit. An initial estimate for the 29 shape parameters is made in the prefit. In the combined fit stage, we would ideally release all these parameters in order to minimize the dependence on Monte Carlo generators. In practice, this is not possible. Every fit parameter increases the chance of instability in the minimization procedure, and thus there is a practical maximum to this number, that can only be found by trial and error. This is the list of 23 shape parameters that combined the largest freedom in shape with a stable fit result.

```
TP_mjjj_peak_mean
TP_mjjj_peak_sigma
CW_mjjj_mean_slope
CW_mjjj_base_sigma
G_mjjj_sigma
CW_mjjj_base_offset
CW_mjjj_mean_offset
CW_mjjj_base_amp
CW_mjjj_base_mean
SU_etmiss_base
SU_mtrans_base
G_mtrans_sigma_ratio
T2_mtrans_sigma_ratio
G_mtrans_corefrac_j4
T2_mtrans_corefrac_j4
CW_mtrans_base
TP_mtrans_base
G_mtrans_sigma
CW_etmiss_mean
G_etmiss_mean
T2_mjjj_base_sig
TP_mjjj_base
T2_mjjj_mean_slope
```

Appendix B

Combined Fit Model on data

The combined fit model used on data in chapter 7 looks like:

$$TT = (1 - f_T 2)TP + f_{T2}T2 \tag{B.1}$$

$$SM = (1 - f_{TT})CW + f_{TT}TT \tag{B.2}$$

$$ALL_{S+B} = (1 - f_{SM})SU + f_{SM}SM \tag{B.3}$$

$$\tag{B.4}$$

where TP denotes the PDF describing the TP background etcetera. Each component model of ALL_{S+B} is a 3D product PDF. Each background model is a 3D conditional product PDF. The composition of these PDFs is given in table B.2 t/m B.3. The naming of PDFS and their parameters is as follows. A 3D PDF is named after the background source it describes. A 1D PDF name contains the source and the observable like < `source`>_< `observable`>. If a PDF is a sum of more than one component, the component PDF names are written as < `source`>_< `observable`>_< `suffix`>, where < `suffix`> will denote the components function like _`peak` or _`tail`. A parameter of a PDF is denoted as < `source`>_< `observable`>_< `parameter`>. This ensures unique names for all PDFs, functions and parameters.

Apart from the tables describing the different models this Appendix also shows for each model the plots that are used to determine the correlations between the different observable. The plots for the CW M_T model can be found in the main text.

PDFs		
PDF	Component	Parameters
TTComb:CW_etmiss		G_etmiss_base, CW_etmiss_mean, CW_etmiss_sigma
TTComb:CW_mtrans		G_mtrans_mean, G_mtrans_sigma, G_mtrans_base
AddPdf:CW_mjjj	Gaussian:CW_mjjj_peak	TP_mjjj_peak_mean, CW_mjjj_peak_sigma, CW_mjjj_fpeak
	TTComb:CW_mjjj_comb	CW_mjjj_base, G_mjjj_mean, G_mjjj_sigma

Table B.1: The CW model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface. The mean of the Gaussian is assumed to be the mean of the TP mass peak, and therefore named **TP_mjjj_peak_mean**, instead of **G_mjjj_peak_mean**.

PDFs		
PDF	Component	Parameters
TTComb:TP_etmiss		G_etmiss_base, G_etmiss_mean, TP_etmiss_sigma
TTComb:TP_mtrans		G_mtrans_mean, G_mtrans_sigma, G_mtrans_base
AddPdf:TP_mjjj	Gaussian:TP_mjjj_peak	TP_mjjj_peak_mean, TP_mjjj_peak_sigma, TP_mjjj_fpeak
	TTComb:TP_mjjj_comb	TP_mjjj_base, TP_mjjj_comb_mean, TP_mjjj_comb_sigma

Table B.2: The TP model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

PDFs		
PDF	Component	Parameters
TTComb:T2_etmiss		G_etmiss_base, G_etmiss_mean, T2_etmiss_sigma
Gaussian:T2_mtrans		T2_mtrans_mean, T2_mtrans_sigma
TTComb:T2_mjjj		T2_mjjj_base, G_mjjj_mean, G_mjjj_sigma

Table B.3: The $T2$ model. PDF parameters that are functions of $E_{\mathrm{T}}^{\mathrm{miss}}$ or other parameters are written in boldface.

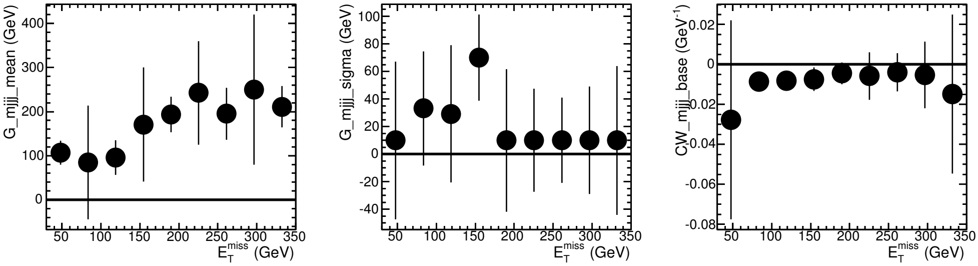


Figure B.1: The fit parameters of the fit of the non-conditional CW M_{jjj} model in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$.

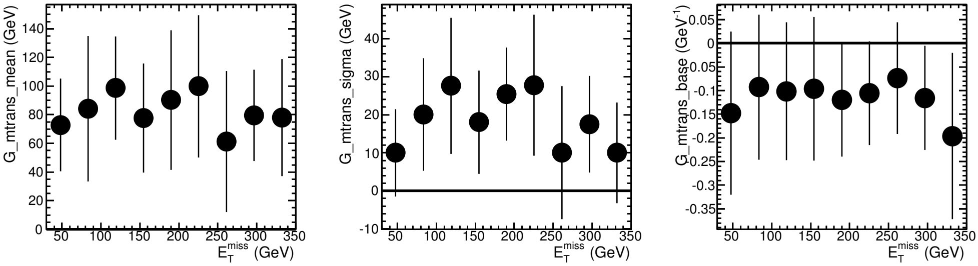


Figure B.2: The fit parameters of the fit of the non-conditional TP M_{T} model in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$.

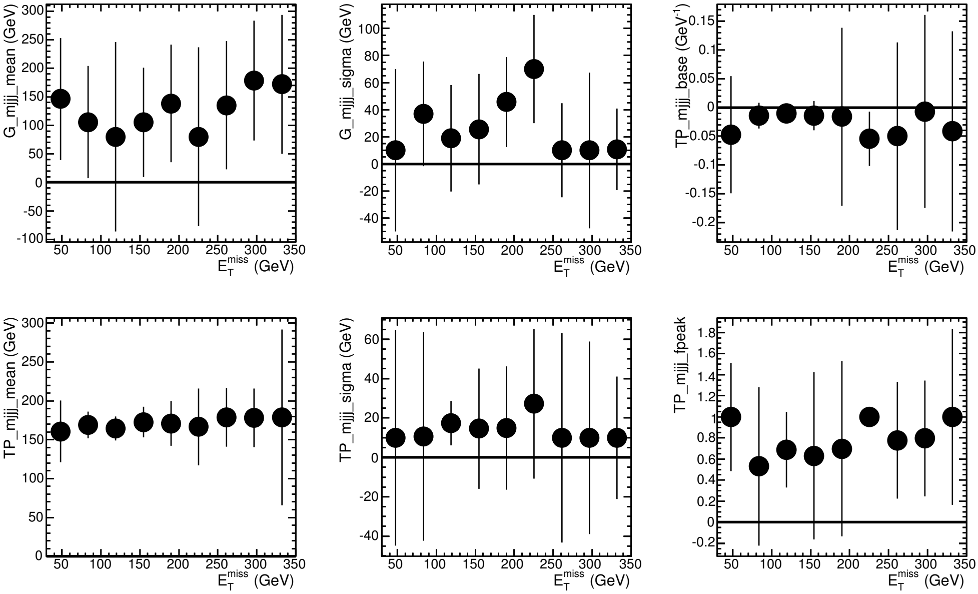


Figure B.3: The fit parameters of the fit of the non-conditional TP M_{jjj} model in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$.

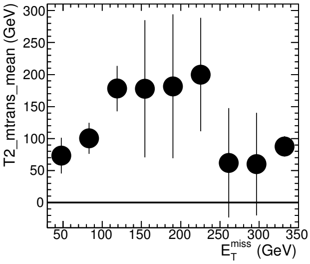


Figure B.4: The fit parameters of the fit of the non-conditional $T2$ M_{T} model in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$.

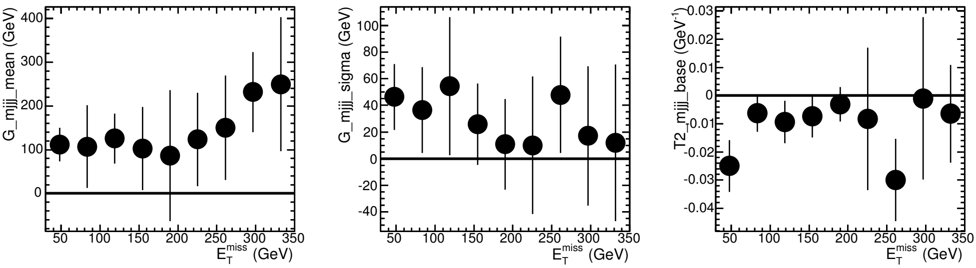


Figure B.5: The fit parameters of the fit of the non-conditional $T2$ M_{jjj} model in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$.

Abstract

In recent years cosmological measurements have shown indications that the particles of the Standard Model (SM) can only account for 4% of the total energy content of the universe. The two other main contributions have been dubbed Dark Energy (74%) and Dark Matter (DM, 22 %). It is fascinating that the theory that is unsurpassed in precision of its predictions by any physical theory turns out to be so limited in scope. Now the remaining 96% of the energy content of the universe can be examined, and a working theory for its makeup can be developed.

One theory predicting a particle that could be the main constituent of DM is Supersymmetry (SUSY). Although no evidence of this theory has been found so far, there is still a large region of the vast SUSY phase space that is unexplored.

The ATLAS detector, one of four experiments on CERN's LHC, is well equipped to do just that. With the data delivered by the LHC during 2010, when two beams of protons were collided at a center of mass energy of 7 TeV, ATLAS can explore a new part of the SUSY phase space, and search for a good candidate DM particle.

Finding SUSY in the ATLAS data is challenging for several reasons. First of all, the precise signature of SUSY in the pp collisions is unknown. Several models exist, each with their own properties and 'smoking guns' that might be present in the ATLAS data. The signature of SUSY that was used in this analysis assumes that whatever supersummetric particle is created by the pp collision, it quickly decays to the lightest supersymetric particle, the LSP, which is assumed to be stable. Such a stable LSP would provide an excellent DM candidate. Since this LSP is subject to the weak interaction only, it will be seen in the detector as missing momentum transverse to the beamline.

The second challenge is to use the missing p_T in a collision as a tool for finding new physics. In order to do this the detector and the objects directly measured and reconstructed (jets, leptons) must be well understood. This is an experimental challenge that was tackled with good results by a large part of the ATLAS collaboration.

The third challenge is distinguishing the SUSY signal from known SM processes. Before 2010, SM processes were not examined at such high energies, and uncertainties in the theory of the SM (e.g. uncertainties on the proton structure, and on finite order cross section calculation) mean that the behavior of the SM at these high energies is not yet well understood. This thesis describes a data-driven method through which the SM behavior is evaluated while simultaneously searching for a SUSY signal. For the two main backgrounds to SUSY, $t\bar{t}$ and W+jets, a model is made of the distribution of missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass M_{T}. The model was built from probability density functions with parameterized shape, which reduced the model dependance on a specific Monte Carlo simulations. Taking advantage of the fact that a key signature of mSUGRA models combines high $E_{\mathrm{T}}^{\mathrm{miss}}$ with high M_{T}, the shape parameters are estimated in a fit to a dataset of events with either low $E_{\mathrm{T}}^{\mathrm{miss}}$ or low M_{T}. With the mass of the three jets with the highest $\sum \vec{p}_T$ (M_{jjj}), a third distribution is added to the model which helps to distinguish the semileptonicaly decaying $t\bar{t}$ background.

When the shape of the model has been estimated in a maximum likelihood fit, the model can be extrapolated to a region with a high signal yield (thus at high $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}), giving an estimate of the number of expected background events in this signal region. Here the shape of the control region, embracing the signal region from two sides with one arm extending to high $E_{\mathrm{T}}^{\mathrm{miss}}$ and one to high M_{T} gives a great benefit to the method in enabling an accurate extrapolation. Although the method works best when the shape of each distribution is independent of the other two observables, in this thesis it will be demonstrated that the correlations between the observables can be accounted for by the model, and dealt with in the fitting and extrapolation procedure.

As the model is normalized to the number of events in the control region, contamination of SUSY events in this region will result in a higher estimate of the number of background events in the signal region, and thus in a reduced sensitivity to new physics. Unfortunately, once the threshold for producing SUSY events is reached, SUSY events will be present also in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ and low M_{T} regions. This thesis will show that it is possible, at least for mSUGRA events, to create a simple Ansatz model for signal in the control region, which can in most cases accurately estimate the level of signal contamination. The main assumption that is made in creating the model is that the mass scale of most mSUGRA models to which the ATLAS detector is sensitive is much higher than for the SM background. The SUSY Ansatz model is not used in the extrapolation.

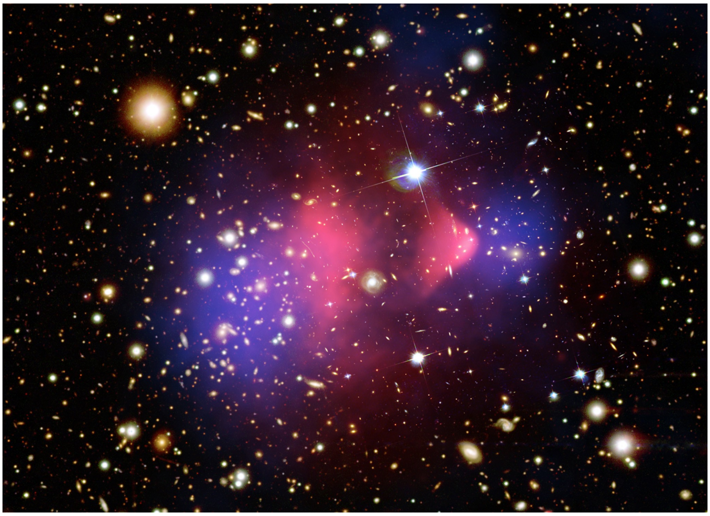
This thesis presents two analyses that were performed using the so-called combined fit method. One analysis is a proof of principle on a relatively large dataset containing simulated events for the SM background, combined with simulated signal events for a wide variety of possible mSUGRA models. A grid was defined in the phase space of mSUGRA parameters, which was used to explore the different behaviors of mSUGRA models. All these samples were created assuming a pp center of mass energy of 10 TeV.

After developing and validating the different tools and aspects of the combined fit method, it was applied on a somewhat smaller (in terms of integrated luminosity) dataset consisting of events measured in the 7 TeV pp collisions recorded by the ATLAS detector during the 2010 data taking period. This resulted in a limit on the number of non-SM events in a dataset with high $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T}, that already exceeds the limit set by previous experiments.

At the time of writing this abstract the 2011 dataset recorded by ATLAS is already 20 times larger than the one used in the final chapter of this thesis. In the analysis of this new dataset the combined fit method can be a valuable tool to find evidence of supersymmetry or other new physics in the ATLAS data, or in the absence of an excess, extend the limit on new physics even further.

Samenvatting

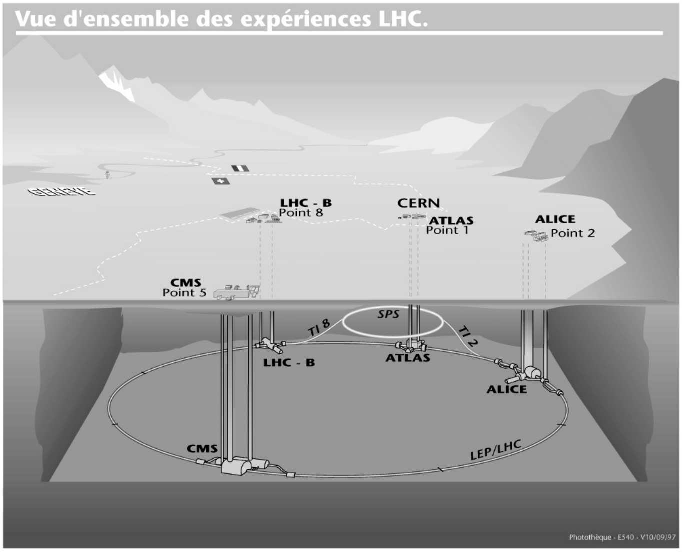
In de afgelopen twintig jaar hebben zowel de deeltjesfysica als de kosmologie grote sprongen gemaakt in het ontdekken en beschrijven van de natuurwetten. Terwijl de deeltjesfysica haar Standaard Model (SM) met steeds grotere precisie kon bevestigen, kwam de kosmologie met steeds beter bewijs dat datzelfde SM slechts een klein deel (4%) van de materie en energie in het universum kan verklaren. Deze 4% bevat de protonen en elektronen en fotonen waar de sterren, planeten en alle dingen om ons heen uit bestaan. De overige 96% bestaat uit Donkere Energie (74%) en Donkere Materie (DM, 22%). In dit proefschrift beschrijf ik een methode om die Donkere Materie nader te onderzoeken met behulp van de ATLAS detector.



Figuur 1: De bullet cluster. De roze vlekken geven de positie van normale materie weer, de paarse vlekken die van Donkere Materie.

De precieze eigenschappen van Donkere Materie (DM) zijn nog niet bekend, omdat nog niemand erin geslaagd is een DM deeltje in het laboratorium te creëren en te bestuderen. Een aantal dingen is echter wel duidelijk. Met het "donker" in donkere materie wordt bedoeld dat de materie geen licht uitzendt, waarbij onder licht alles wordt verstaan van radiogolven tot gamma straling. Dat betekent dat een donker deeltje geen elektrische lading kan hebben.

Verder is bekend dat DM slechts zeer zwak met zichzelf en zijn omgeving reageert: in Figuur 1 zijn twee melkwegstelsel te zien die met elkaar in botsing zijn gekomen. De roze vlekken op de foto geven de positie van gaswolken en sterren weer die door de botsing zijn



Figuur 2: De tunnel van de LHC versneller, tussen de Jura en het vliegveld van Genève. De 4 locaties waar protonbundels botsen zijn in dit plaatje aangegeven.

afgeremd. De paarse vlekken geven de locaties weer waar de DM zich moet bevinden. Dit is af te leiden uit de zwaartekracht die door de DM wordt gegenereerd, en die het licht van achterliggende sterrenstelsels afbuigt. Op de foto is te zien dat de DM niets gemerkt heeft van de botsing, maar schijnbaar ongehinderd is doorgevlogen.

Als een DM deeltje niet of nauwelijks reageert met zichzelf of zijn omgeving, dan reageert het ook niet met het materiaal waar de detectoren van ATLAS van gemaakt zijn. Een DM deeltje is dus hoogstwaarschijnlijk onzichtbaar voor deze detectoren. Dit maakt het een uitdaging om aan te tonen dat dit deeltje in het lab is gecreëerd. Die uitdaging ben ik in dit proefschrift aangegaan.

Uit eerdere pogingen om een DM-deeltje te maken en te detecteren is gebleken dat dit deeltje erg zwaar moet zijn, anders was het inmiddels gevonden. De massa van dit nieuwe deeltje moet minstens 100 keer zo groot zijn als de massa van het proton. Om het zware deeltje te maken worden twee bundels protonen met hoge energie rondgeleid door de LHC, de Large Hadron Collider. Als de twee bundels op elkaar worden gericht, kan proton uit de ene bundel botsen met een proton uit de andere. Bij die botsing wordt een enorme hoeveelheid energie samengebracht in een punt in de ruimte. Die energie kan worden omgezet in een nieuw deeltje. Dit deeltje zal in de regel heel snel uiteenvallen in een groot scala aan andere deeltjes: elektronen, muonen, quarks, en soms ook, als het bestaat, een DM-deeltje.

Die botsingen vinden plaats op vier locaties in de ring van de LHC versneller. In Figuur 2 is de tunnel van de LHC versneller te zien, 27 km lang, 100 m onder de grond. Op de vier punten waar botsingen plaatsvinden zijn detectoren gebouwd die het resultaat van die botsingen meten. Op één van die vier locaties bevindt zich de ATLAS detector. De protonen botsen op elkaar in het binnenste van de ATLAS detector. De impuls, massa en lading van de vervalsproducten van de botsing worden gemeten door de detector, en die metingen worden gebruikt om te herleiden wat er in de botsing precies gebeurd is.

Aangezien het DM deeltje onzichtbaar is, moet er een truc worden uitgehaald om het zichtbaar te maken. Van alle zichtbare deeltjes die uit een botsing komen wordt de impuls bij elkaar opgeteld. Hierbij wordt alleen de impuls loodrecht op de richting van de protonbundels meegerekend. Aangezien de totale impuls altijd behouden is (volgens de wetten van Newton), en aangezien de impuls van de protonen loodrecht op de bundel gelijk aan nul is, moet de totale impuls van de vervalsproducten ook gelijk zijn aan nul.

Als nu een van die vervalsproducten onzichtbaar is, wordt die impuls niet meegenomen in de berekening. Die is namelijk niet gemeten. De totale impuls van alle gemeten deeltjes is dan ongelijk aan nul. Daarmee is de impuls van het onzichtbare deeltje zichtbaar gemaakt. Hier dient zich de eerste uitdaging aan: als de detector niet helemaal naar behoren functioneert, en de impuls van een deeltje wordt niet goed gemeten, dan wordt de balans verstoord. De meting suggereert dan onterecht een onzichtbaar deeltje. Als de impuls balans gebruikt wordt in een meting moet de detector goed werken, of in ieder geval moeten de tekortkomingen van de detector goed begrepen zijn, zodat ervoor gecorrigeerd kan worden.

Als nu het deeltje van de Donkere Materie het enige onzichtbare deeltje zou zijn, zou dit onderzoek redelijk eenvoudig zijn geweest. Dit is echter niet het geval. In het Standaard Model zitten ook enkele onzichtbare deeltjes, de neutrino's. Om te bewijzen dat onzichtbare impuls kan worden toegeschreven aan een nieuw deeltje, en niet aan neutrino's, moet eerst de onzichtbare impuls van de neutrino's goed gemeten zijn.

Hier stuit men weer op een nieuwe uitdaging. Hoewel de deeltjes van het SM model goed bekend en onderzocht zijn, zijn ze nog nooit eerder gecreëerd in botsingen met zo veel energie als bij de LHC. Simulaties van botsingen waarbij SM deeltjes worden gemaakt zijn dan ook om verschillende reden onbetrouwbaar. Daarom was de eerste stap van deze analyse het maken van een model van de productie van de SM deeltjes, waarbij de onzekerheid in het model gestopt is door middel van parameters die de vorm van het model beschrijven. De waarde van die parameters kan worden vastgesteld door het model met de data te vergelijken. Hiermee is het model flexibeler gemaakt, waardoor het de data ook goed kan beschrijven als de botsingen van protonen er in de detector iets anders uit zien dan wat men op basis van de simulaties zou verwachten.

Het zijn echter uiteindelijk niet de deeltjes van het SM die interessant zijn voor mijn onderzoek. Naast de achtergrond van bekende processen uit het Standaard Model, moeten ook de nieuwe deeltjes door het model beschreven worden. Om een idee te krijgen hoe het signaal van DM gemeten door de ATLAS detector eruit zouden komen te zien heb ik een theorie gebruikt die een deeltje voorspelt dat alle eigenschappen van een DM deeltje bezit: het is onzichtbaar, stabiel, en kan zwaar genoeg zijn. Deze theorie wordt supersymmetrie genoemd. De symmetrie van deze theorie voorspelt een nieuw "spiegelbeeld" deeltje voor ieder deeltje van het SM. Daarnaast wordt er nog een aantal nieuwe deeltjes voorspeld. Een van de deeltjes die ingebed zitten in de theorie van supersymmetrie is de zogenaamde neutralino. Dit is een goede kandidaat voor een DM deeltje.

Er zijn vele verschillende varianten van supersymmetrische theorieën. Zelfs als men zich beperkt tot één variant, zoals ik in dit proefschrift heb gedaan, dan zijn er vele verschillende verschijningsvormen afhankelijk van de parameters van de theorie.

Op basis van vele simulaties van verschillende vormen van deze theorie heb ik een model gemaakt dat voorspelt hoe de productie en het verval van supersymmetrische deeltjes in de ATLAS detector eruit zou komen te zien. De twee modellen, van de SM achtergrond en het signaal van supersymmetrie heb ik gecombineerd. Door dit model (genaamd "the combined fit model") te vergelijken met de data heb ik kunnen aantonen dat voor een groot aantal supersymmetrische modellen geen bewijs is gevonden in de ATLAS detector. Daarmee kunnen die modellen uitgesloten worden voor verder onderzoek.

Dit resultaat is verkregen met een hele kleine dataset: de hoeveelheid data die inmiddels is verzameld is 30 keer groter dan wat er voor dit proefschrift beschikbaar was. Zelfs met deze kleine dataset heb ik al een groot aantal verschijningsvormen van supersymmetrie kunnen uit sluiten. Met de nieuwe data van ATLAS kan een veel groter scala aan theorieën getest worden. Het is heel goed mogelijk dat daarbij bewijs wordt gevonden voor het DM deeltje waar ik naar op zoek ben geweest. De methode die in dit proefschrift is ontwikkeld zal een substantiële bijdrage leveren aan deze zoektocht naar nieuwe fysica.

Acknowledgements

Finally, it is time to write the most important part of this thesis. Between typing the first little script that would eventually lead to an analysis, and writing these acknowledgements, many things have happened. I have met a great number people, worked with some of them, drank, skied, dined, discussed with many others, married one. Now is the time to thank all of them. Some, I want to name specifically.

First, my promotores. Nicolo de Groot I thank for unfailing confidence in the success of this project, and insightful discussions on the nature of physics and (related) politics. Paul de Jong I thank for the detailed critical comments that were crucial, particularly in the final stages of writing this thesis.

Wouter Verkerke has been an excellent guide through the swamps of statistics, programming and RooFit, although it must be mentioned that the latter was a swamp of his own creation. Marcel Merk has kept a watchful eye on my progress, and has enabled me to finish my thesis on a chapter with the first ATLAS data. Max Baak I thank for patiently and carefully explaining the SUSY group methods of limit setting. Peter Jansweijer has been an inspired and inspirational supervisor in my brief encounter with electronic engineering. Irene Niessen, thank you for being a constant companion and fantastic friend from the beginning of my studies in High Energy Physics. Moreover, you were an invaluable source of knowledge and understanding when writing the first chapter of this thesis on Supersymmetry theory.

Alex Koutsman has shared all the frustrations and successes which were part of doing this (or any) analysis, and with him I share the pride of seeing the combined fit method complete, and of knowing that it will be a stepping stone to discovery (or exclusion) of SUSY.

My office mates in old and new buildings, in the Netherlands and Geneva: Eric, Lucian, Melvin, Geert-Jan, Marcel, Michele, Miruna, Raoul, who have felt the pain of many keyboards, and worked to the background noise of constant swearing, I thank for the diversions which often had nothing to do with work.

Kees de Lange, Michel Bochem, Cor Heesbeen en Connie Morsing hebben mij veel geleerd over les geven, en hebben mij de enorme voldoening van een geslaagde les laten meemaken. José Coppens dank ik voor de fijne samenwerking, en de reflecties op het vak van leraar.

De sfeer op een afdeling is uiteraard ontzettend belangrijk voor het plezier in het werk. Marjo, Gemma and Annelies, dank voor de gezelligheid. Sinterklaas, dagje uit, borrels, maar vooral ook de koffiepauzes in jullie kantoor.

Els, Lex, Krijn, Daniela. Waar te beginnen. Ik noem als voorbeeld voor de cruciale momenten waarop jullie er voor me waren een terrasje aan de Waal aan het begin, en een zolderkamer waar ik mij op kon sluiten helemaal aan het eind van mijn proefschrift. Er waren er veel, veel meer. Zonder jullie was dit proefschrift er niet geweest.

Mijn broers en zussen, Yulianto, Joris, Michiel, Brenda, Tessa, Joost, Giulia, Marianna en mijn schoonfamilie Bert, Tiny, Susan en Michel, ik heb jullie de oren van de kop geluld met verhalen over protonen, electronen en donkere materie. Dank voor het luisteren, dank voor het vragen stellen. En wees niet bang, ook over radiotherapie kan ik vast nog wel een hele avond ouwehoeren... Luuk, Timo en Mees, jullie waren te jong om mee te praten over mijn werk, maar ik wil jullie noemen, omdat ik altijd vrolijk word als ik jullie zie. Michiel wil ik ook bedanken voor het ontwerp van de fantastische cover van dit proefschrift.

Opa Stomphorst dank ik voor het maken van de illustratie op de cover van dit proefschrift, maar vooral voor alle getoonde interesse in de vorm van krantenknipsels, verhalen, vragen en discussies.

Mijn paranimfen, Lennart en Robert-Jan. Al heel lang maken jullie alle belangrijke stappen in mijn leven van dichtbij mee. Ik ben heel blij om jullie ook in het laatste uur van mijn promotie achter mij te weten.

Mijn laatste dank is aan mijn vrouw, Inge. Dankjewel voor de ontelbare keren dat je me gesteund hebt als ik het even niet zag zitten, en voor de minstens even vaak noodzakelijke schop onder mijn kont als ik dacht dat ik mijn werk wel kon ontvluchten. Je hebt me inspiratie, zelfvertrouwen en doorzettingsvermogen gegeven die ik nodig had om dit tot een goed einde te brengen. Nu, twee proefschriften later, weten we wat we aan elkaar hebben in tijden van stress. Dat komt goed uit; op naar het volgende avontuur.

Curriculum Vitae

Folkert Koetsveld was born on January 10$^{\text{th}}$, 1979 in the Dutch city of The Hague. He studied Astronomy at the University of Utrecht, where he obtained his Master's degree in 2002. He completed a second specialization in 2006, obtaining a Master's degree in high energy physics at the Radboud University Nijmegen. During this specialization he spent three months at the CERN Summer Student programme in Geneva, where he was involved in the calibration of the ATLAS calorimeter.

In 2006 he started the PhD project on a data driven method for Supersymmetry searches at the ATLAS detector, which resulted in this thesis. For this project he worked at the CERN international laboratory for subatomic research for one year. Since August 2011, he works at the Radiotherapy department of the NKI-AVL on image guided adaptive radiotherapy (IGART) techniques.